TURING 图灵新知

The Great Equations

Breakthroughs in Science from Pythagoras to Heisenberg

历史上最伟大的10个方程

[美] Robert P. Crease 著
马潇潇 译



$F_g = \frac{Gm_1m_2}{r^2}$

$F=ma$

$e^{i\pi}+1=0$

$E=mc^2$

$c^2=a^2+b^2$



人民邮电出版社
POSTS & TELECOM PRESS

图书在版编目（C I P）数据

历史上最伟大的10个方程 / (美) 克里斯(Crease,R.P.)
著 ; 马潇潇译. -- 北京 : 人民邮电出版社, 2010.8
（图灵新知）
书名原文: The Great Equations: Breakthroughs
in Science from Pythagoras to Heisenberg
ISBN 978-7-115-23175-8

Ⅰ. ①历… Ⅱ. ①克… ②马… Ⅲ. ①方程－普及读
物 Ⅳ. ①O122.2-49

中国版本图书馆CIP数据核字(2010)第122127号

内 容 提 要

本书按方程出现的时间顺序，依次介绍了毕达哥拉斯定理、欧拉公式、麦克斯韦方程组和海森堡不确定性原理等 10 个历史上最伟大的方程。作者以科学历史学家和哲学家的双重视角，采用跨学科式的写作技巧，告诉读者任何一个重要方程的出现都不是哪个天才一时兴起想出的，而是经过数年、数十年甚至数百年的文化发展才达到的顶峰。他还告诉人们，方程所影响的不仅是科学和数学，它们甚至会改变人们的思维方式。

本书适合对科学史、数学和物理学感兴趣的各层次读者。

图灵新知

历史上最伟大的10个方程

◆ 著 [美] Robert P. Crease
译 马潇潇
责任编辑 马晓燕

◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街14号
邮编 100061 电子函件 315@ptpress.com.cn
网址 http://www.ptpress.com.cn
北京 印刷

◆ 开本：880×1230 1/32
印张：8.25
字数：221千字 2010年 8月第 1 版
印数：1－4 000册 2010年 8 月北京第 1 次印刷

著作权合同登记号 图字：01-2009-6935号

ISBN 978-7-115-23175-8

定价：29.00元

读者服务热线：(010)51095186 印装质量热线：(010)67129223

反盗版热线：(010)67171154

版 权 声 明

关于作者

Robert P. Crease 是纽约石溪大学哲学系教授和系主任，同时也是布鲁克海文国家实验室的历史学者。他一直为《物理世界》杂志的每月专栏“Critical Point”供稿。他的著作有《棱镜与钟摆：10 个最漂亮的科学实验》（*The Prism And The Pendulum: The Ten Most Beautiful* Experiments in Science），《创造物理科学：布鲁克海文国家实验室的历史》（*Making Physics: A Biography of Brookhaven National Laboratory*），《自然之剧：实验还是表演？》（*The Play of Nature: Experimentation as Performance*），《再创世：21 世纪物理学革命的缔造者》（*The Second Creation: Makers of the Revolution in Twentieth-Century Physics*, 与 Charles C. Mann 合著），《战争与和平：回忆科学前沿的日子》（*Peace & War: Reminiscences of a Life on the Frontiers of Science*，与 Robert Serber 合著）。Crease 的译作有《美国技术哲学：经验转向》（*American Philosophy of Technology: The Empirical Turn*）和《技术、代理和设计的哲学思考》（*What Things Do: Philosophical Reflections on Technology, Agency, and Design*）。他经常应邀作报告，有许多文章和评论发表在《大西洋月刊》、《纽约时报》、《华尔街日报》、《科学》、《新科学家》、《科学美国人》和《史密森尼》等杂志上。作者现居纽约市。

致　谢

与前一本书《棱镜与钟摆》（*The Prism and the Pendulum*: *The Ten Most Beautiful Experiments in Science*）一样，本书也是在本人为《物理世界》所写专栏的基础上写成的。在此，我要再次对本书的诸位编辑表示感谢，特别要感谢让我有机会为《物理世界》撰写专栏的 Matin Durrani，以及对专栏中的伟大方程提出意见的上千位读者。得益于这一专栏，我才能在书中试着去提出各种思想。读者也会发现，专栏中的内容在书中通篇可见。我还要感谢版权代理 John Michel 不厌其烦地为我指明正确的方向，以及在本书写作过程中给予帮助的 Margaret Maloney、产品经理 Julia Druskin 和编辑 Maria Guarnaschelli。Maria 通读了本书，并提出了指导意见。与其他专栏作家一样，我从同事和通信联系人那里也受益匪浅，他们给了我灵感、思想和信息。这些向我提出有用建议和意见，并给予各种帮助的人有：Edward S. Casey、David Cassidy、Carlo Cercignani、Allegra de Laurentiis、John de Pillis、David Dilworth、B. Jeffrey Edwards、Elizabeth Garber、Patrick Grim、Richard Harrison、George W. Hart、Richard Howard、Don Ihde、Eric Jones、Ed Leibowitz、Gerald M. Lucas、Bob Lloyd、Peter Manchester、Eduardo Mendieta、Hal Metcalf、Lee Miller、Eli Maor、Anthony Phillips、Xi Ping、Mary Rawlinson、Robert C. Scharff、David Socher、Marshall Spector、Clifford Swartz、Dick Teresi 和 Beth Young。如果没有哲学系办公室的 Alissa Betz、AnnMarie Monaghan 和 Nathan Leoce-Schappin 的帮助，本书不可能这么快就能与读者见面。在关于量子力学的几章中，John H. Marburger, III 帮我避免了许多理解上的错误。其实，要感谢的人还有许多。Alfred S. Goldhaber 就有关量子力学和相对论的章节提出了鞭辟入里的建议。我们二人曾合教过一门介绍量子力学

在物理学领域之外的影响的课，在同他的讨论中我同样获益良多。我的妻子 Stephanie 不仅通读了初稿，一直以来还默默忍受着我的繁重工作。这对我的家人来说是严峻的考验，但她总会带给我惊喜。我的儿子 Alexander 也是默默忍受着我的工作习惯。虽然时常见不到我，可他毫无怨言。我的女儿 India 总能让我行进在正确的方向上。在别人都不理它的时候，我的小狗 Kendall 老是希望我能带它出去走走。最后，我还要感谢当今最出色的科学作家 Charles C. Mann。他是我的榜样，感谢他的慷慨和给予我的灵感。

引　言

多数人学到的第一个方程[①]便是 1+1=2。它虽然简单得不能再简单，却有巨大的影响力。它给出了加法的定义：1 加 1 等于 2。它有用的另一原因是，算术、整个数学领域、物理学和其他科学分支领域中的所有方程，都是这种形式。方程把各项进行合理的安排，指出这些项之间的某种特定关系。1+1=2 这个方程很小、很简单，也很基本，但它却像一把魔杖，开启了许许多多领域的大门。1+1=2 是开启知识之门的小小第一步，却也是后面成千上万步的基础。加拿大卡尔加里皇家山学院的诗人和英语教师理查德·哈里森（Richard Harrison）在给我的信中曾提到这一表达式的深远影响。

> 1+1=2 是数学中的童话，是我教儿子的第一个方程，也是揭示人类具有改变现实世界的神奇力量的第一个表达式。我还记得儿子在学习这个表达式时伸出两只手的食指——代表 1，脸上满是惊奇的表情。他发现被自己身体分开的两个手指，通过头脑中的一个概念就能联系起来。这或许是他第一次真正意义上的哲学探索。当发现儿子的思维能够理解“1+1”所代表的不仅仅是“1+1”的时候，我把这个小小的方程视为儿子认识自我和他人而不是神奇外部世界的金钥匙。

哈里森的描述提醒我们，即使是像 1+1 这样非常基本的方程，学习的过程实际上也是一段“旅程”。这段“旅程”包括三个阶段：一开始我们很幼稚，不知道方程的存在；通过学校的教育、偶然的习得、好奇心

① 严格来说，1+1=2 并不是方程，只是一个等式，此处是为了与书名呼应。本书的主题讲的是“equation ”，它既指方程，也指等式、公式。——编者注

或是有目的的学习，我们开始逐渐理解方程，这期间少不了不满足和挫折感；最后，我们对于世界的体验因为学习的经历发生转变，变得对世界充满好奇，这种转变自然无痕，虽然有时候只是暂时的。

本书要讲的就是这些“旅程”。

人们最初并不知道方程，它们好像也没有存在的必要。不管是在“伊甸园”里，还是“知识树”上，都找不到方程的影子。无论是苏美尔人的天堂迪尔蒙（Dilmun）里，还是中国传说中盘古开天劈地的宇宙大蛋里，亦或是其他各种创世神话中所描述的人类最早出现的地方，都看不到方程的影子。那时候的人没有方程的概念。方程的思想是人发明出来的，是人类努力认识世界的结果。即便如此，方程也绝不是人们哪天早上醒来后一时兴起就决定发明的。随着时间的流逝，人们有了对于方程的需求，真正科学技术意义上的方程概念是出现在人类历史较晚的时候。

拉丁词 aequare 的意思是保持平衡或者相等。现在的许多英语词汇都是从这个词根衍生出来的，包括 adequate（合适的）、equanimity（平静）、equality（平等）、equilibrium（平衡）、egalitarian（平等主义）、equivalence（相等）和 equivocation（模棱两可）。“方程”（equation）一词本身的含义很简单：就是分成相等的两份。再如，地理学家为将地球分成大致相等两半，想象出来的一条线叫做赤道（equator）。equation 一词被中世纪的占星家用来说明沿着太阳和其他行星的轨道将天空分成面积相等的块的做法，每块分别由一个星座掌管[1]。

与此同时，数字和计数法在人们生活中也变得日益重要。商人将数字和计数法用于记账、财务和预算，宗教机构用它们记录年代、季节和生死嫁娶等，而政府官员将其用于人口普查、土地测量和税收[2]。这样就产生了发明符号来表示数字和数量的需求[3]。公元 3 世纪，希腊数学家丢番图更进一步，他采用符号代表*未知量*，并给出了操作这些量的方法，包括加减。他不仅说明了如何用符号表示未知数，以便通过已知

数将其求出（对应于定方程），还指出符号可以描述无穷多解的情形（对应于丢番图方程，或称不定方程）。不过这与现代的方程思想仍相去甚远。甚至连伽利略和牛顿在表述他们各自的重要结果，即伽利略自由落体定律和牛顿运动定律时，也仍旧采用文字叙述的方式来表示比例，而不是采用如今理科生都很熟悉的方程。直到 18 世纪，自然科学家们才普遍采用我们今天所熟知的方程形式来表述他们的结论。

因此，即使要写出最简单的方程，也得经历一番概念形成的发展历程。1910 年，历史上两位著名的数学家，阿尔弗莱德・诺斯・怀特海（Alfred North Whitehead）和伯特兰・罗素（Bertrand Russell）出版了著名的三卷系统教科书《数学原理》（*Principia Mathematica*）。该书以纯逻辑方式从头推导出了数学的基础。即便要找到 1+1=2 这样一个简单的方程，也得翻到第一卷的后半部分去[4]！

经过这段漫长的历程，"方程"一词最终才有了专业的意义，成为特殊人造语言的一部分。用它可以描述两个相同的量，或者两个相同的可测量集合（严格地说，表述不相等关系的不等式不能算作方程）。这种像密码一样的人造语言对于现代数学和科学是不可或缺的。这种语言的符号代表了其他事物的集合。我们可以对这些符号进行各种操作，最简单的是加减乘除。[5]

在这种特殊的专业语言发明之后，每个方程都具有两种不同类型的变形。它是由最初将方程引入人类文明的人发现的。自此以后，所有学习过方程的人都会发现这一点。

方程的发现背景与历史上其他转折点的发生背景不同。在血腥的战场和激烈的政治冲突中是不会有方程出现的。它的出现需要不受世事纷扰的安静环境，比如书房或图书馆。麦克斯韦在书房中写出了他革命性的方程；海森堡则在孤岛上将点滴碎片统一了起来。科学家在这样的环境里可以设法消除不满足感，对付痛苦。这种痛苦的出现是因为手头的

点滴碎片不能很好地统一起来，而需要作出调整或者加入新元素。如此一来，科学家就可以把精力集中到某些描述得很简单，其实迷惑性却很强的问题上。例如，直角三角形某条边的长度是多少？天体之间引力的大小是多少？电是如何流动的？一对看似矛盾的理论可以被统一起来吗？这些有意义吗？

只要答案一出现，它似乎就是符合逻辑的，甚至是必然的。罗杰·科茨（Roger Cotes）曾在牛顿的大作《自然哲学的数学原理》（*Mathematical Principles of Natural Philosophy*）第 2 版的前言中写道，该工作“得到了广泛的认同”[6]。发现者往往是偶然发现了业已存在的事实。这些方程就像宝藏，在尚未成形时被有洞察力的科学家发现。科学家随后对它进行研究，最后放到知识的大宝库中去，从此代代相传。这种方法对于呈现科学发现极为方便，多为教科书所采用，不妨称为宇宙藏宝图。这张图为我们一一展现了发现方程的艰难历程，以及方程的发明者、发现时间和地点，通常还有起因和目的等。有时候，一个很长的发现历程经过提炼后，会浓缩成某个事件，例如苹果的下落。历代学者在批评这种模式和完善这张图的过程中赚足了名声。看来“藏宝图”真是对人人都有益！

但是不管宇宙藏宝图多么有用，这些方程也不过是对宇宙基本特征的描述，而不是人类的创造。的确，我们降生到这个世界之前，这些方程就已经存在了。这也正是方程有时看似并非由人力产生的原因。据说，方程很早以前就存在了，创世第八天，上帝创造了方程，以它们作为工作的蓝图。正如伽利略所写的那样，“自然之书”就是用数学符号写成的。

但是所有的方程又都有人的参与。每个方程都是由特定的人，在特定的时间和地点提出来的。他感到有必要对手头上不令人满意的资料进行解释，不过有时可能仅是为了使过于复杂的知识变得更容易理解一些。这种创新的过程有时候是隐含在过去的资料中的，例如毕达哥拉斯定理①

① 中国称作勾股定理。——编者注

早在毕达哥拉斯之前很早就已为人所知。还有些时候，多亏有方程发明者的信件、手稿和笔记，我们才能了解到创新过程的细节，牛顿和爱因斯坦提出的方程就是如此。但不管是哪一种情况，任何一个方程的发现，都不能归功于某一个人。因为哪怕科学家是一个人独立工作的，他也需要与同行进行充分的对话交流，来共同理解和解释世界。

当英国科学家奥利弗·赫维赛德（Oliver Heaviside）把麦克斯韦的工作整理成现在人们所熟知的形式——现在所谓的“麦克斯韦方程”时，他说自己只是想要试着更清楚地理解麦克斯韦的工作。这种动机——想要更好地表述已知的，但并不是特别明白的东西，可以说是所有方程提出者的共性。

在有人针对基本问题提出新方程，解决了自己之前的不满意之处后，人们和世界也会随之改变。这些方程并非只是告诉人们如何去做计算，或者仅仅是作为新增的工具，它们有着更加深远的意义。就像哈里森所指出的那样，他的儿子学习 1+1=2 之后，并非仅仅输入一个新数据点，而是通过这种学习，孩子对世界有了更加深入的把握。当然，伴随着这种把握而来的还有迷惑和新的不满足[7]。

最后，哈里森告诉人们方程可以激发好奇心。科学研究不是我们冷眼看世界的活动，而是一种很微妙的情感生活方式。获得新发现或取得新成绩，谁都会高兴雀跃，加以庆祝，这是很自然的。但如果做科研的人只想着高兴，把科学发现作为获取功名的手段，那就大错特错了。功成名就不仅人数寥寥可数，而且实现起来也是遥遥无期。幸运的是，科学所涉及的情感要广泛而深厚得多。科学研究的过程就是时时刻刻表露情感的过程——疑虑、困惑、好奇、渴望和想要找到答案的迫切心情，进展停滞时的厌倦情绪和挫败感，以及进入正轨后的激动心情。这样的情绪是时刻存在着的，又并非隐藏在内心深处，常常会被疏忽，不过只要稍加留意就会发现它们的存在。

人们一旦理解了某个重要方程，就能窥见比我们感知的更深层面上的世界结构，揭示世界本身与人们经验之间的深层联系。这种时刻，人们的反应并不仅仅是“对，可以理解了”，更常听到的则是“啊哈，原来如此”。后一种是伴随着知识藏宝图的获得而来的，它把获得发现后的情感简化和浓缩成了一个瞬间。而真正的情感——好奇心，则更加细腻、丰富和持久。

当科学家专注于世界，对世界产生兴趣时，他们会很自然地失去对于方程的好奇心，不再那么注意当初公布发现的方程时的情景。的确，人们对于自己非常熟悉的仪器或事物常会失去好奇心。方程可能被看作是人们了解和发现世界的工具，也可能被看作是不得不学的烦琐事物。

马克·吐温（Mark Twain）在《密西西比河上的生活》（*Life on the Mississippi*）一书中写道，领航员对船越来越熟悉，却常常需要经历一种令人遗憾的转变。随着对河流解读技术的增长，他们慢慢地不再感受得到河流的美丽和诗情画意。漂浮着木筏的河面，水面上的倾斜标记，还有一片片的水纹，曾一度激起他们的好奇和敬畏。如今，他们对河的欣赏却越来越机械，河的用处好像只是领航。人们看待方程的态度也是这样。

不过大科学家依旧能够对前人的重大发现保持好奇心。物理学家弗兰克·维尔泽克（Frank Wilczek）曾写了一系列的文章，用简单方程对牛顿第二运动定律 $F=ma$ 进行表述。他将该定律称为“经典力学的灵魂”，并用一种人们容易理解和接受的方式去呈现它[8]。物理学家和宇宙学家钱德拉塞卡尔（Subrahmanyan Chandrasekhar）曾针对牛顿的《自然哲学的数学原理》专门写了一本书（牛顿在《自然哲学的数学原理》中提出了他的第二运动定律），认为该书可与西斯廷教堂天顶的米开朗基罗画相媲美。理查德·费曼的《物理学讲义》的一位听众发现，费曼在教学生方程的讲解中充满无所畏惧的、主动的好奇心。这三位诺贝尔奖获得者非常清楚应该如何保持对于世界和方程的好奇心。好奇心是了解和认识世界的大门。

本书旨在说明方程的深刻含义，告诉读者它们不只是简单的工具。方程与其他人类智慧的结晶一样，有着重要的社会意义，可产生文化力量。本书选取了一些重要方程，扼要介绍了它们的发现者、发现的缘起以及它们怎样解读世界的本质。

注　　释

[1] 有一个事实是，星座并不是固定不变的。太阳和恒星在经过星系的时候，在不同的时间有时经过的是完全不同的星系，而并非报纸上的黄道十二宫图宣称的日期。但现代天文学家似乎并没有受到什么影响，甚至对此完全无视。看来需要有人撰文指出他们的失职了。

[2] 参见伯纳德•科恩的《数字的胜利：计数法如何塑造现代生活》（*The Triumph of Numbers:How Counting Shaped Modern Life*）（纽约：诺顿出版公司, 2005）

[3] 这样一来就得采用符号表示数字和其他事物。古埃及人、巴比伦人和希腊人分别发明了不同的数字和数量表示法。当时的很多古代数学家就是通过研究特定情形的解，然后让读者将其一般化的。例如，公元前 1650 年著名的埃及手稿——《林德手卷》（*Rhind papyrus*）中就有了基本方程的雏形。这些雏形都是基于实例的，计算了三角形、矩形和圆形的面积，以及三棱锥和圆柱体的体积。手稿还给出了许多实际问题的解，其中有如何确定不同密实度的面包块与不同量的大麦之间的等量关系。手稿甚至还讨论了没有实际意义，但在概念上却很重要的代表性问题，例如这样一个问题：“有 7 间房子，每间房子里有 7 只猫，每只猫能吃 7 只老鼠，每只老鼠吃了 7 粒大麦，每粒大麦又能产出 7‘hekat’①大麦。问上面列出的所有都加起来是多少？”此后又出现了许多该问题的不同版本，如古斯大妈的“来自圣艾夫斯的人”。这个人有 7 个老婆，每个老婆有 7 个袋子，每个袋子里有 7 只猫，每只猫又生出 7 只小猫。希腊人在方程中采用象形文字作为符号。象形文字就像人的两条腿一样。做加法时，“两条腿”沿着书写的方向走；做减法时，“两条腿”沿着相反的方向走。

[4] 参见怀特海德和罗素，《数学原理》，第 1 卷，（剑桥：剑桥大学出版社，1957），第 362 页：“从该命题出发，就能定义算术的加法，即 1+1=2。”

---

①古埃及体积计量单位。——译者注

[5] 如今有几种不同的方程分类法。一种是按方程的次数分类，也就是按最大的指数项分类。线性方程之所以被称为“线性”，是因为它描述的是线（如 $4x+3y=11$ 和 $y=2x+1$）。因为未知数 x 和 y 上没有标明次数，所以都是一次方的。如果把未知数平方，则方程就变成二次方的；如果把未知数立方，则方程就变成三次方的；以此类推就能得到四次方、五次方和六次方的方程。如果方程的解不是数字，而是包含微商的函数，该方程称为微分方程。

[6] 参见牛顿的《自然哲学的数学原理》一书，伯纳德·科恩和安妮·惠特曼译（伯克利：加利福尼亚大学出版社，1999 年），第 391 页。

[7] 人们由此出发比较了方程和诗。这两者对于语言的运用都超越了常人的水平，以期能够简明准确地传达出真实来。这样一来，原本难以理解的东西，就变得清楚了。经过这个过程，人们的体验感受会发生变化。迈克尔·季伦（Michael Guillen）在 *Five Equations that Changed the World: The Power and Poetry of Mathematics* 一书中写道，方程“以独特的精确性描述事实，以极其简明的语言传达出大量的信息。外行一般很难理解它的”。季伦还写道：“传统诗歌让人们看清自己的内心，数学之诗则让人们的认识超越自我。”《天地有大美》（*It Must Be Beautiful: Great Equations of Modern Science*）一书的编辑格雷姆·法米罗（Graham Farmelo）也把方程与诗歌做了比较。他认为，虽然两者中有很多单独项是没有特指的，但它们都是人们在认识世界的过程中抽象出来的。法米罗还写道，“诗歌是最精确语言的高度浓缩”，而方程是“在物理现实层面理解世界的最简洁形式”。

当然，诗歌和方程之间也有诸多差异。方程看起来更可怖一些。它们不仅超出了人们的理解能力，而且涉及的是人力不可控的力量，令人感到无助和愤恨。诗歌一般涉及的是人对于周围世界的直观感受。它通过直接影响人的直观感受唤起人的情感，而不提供信息。相反，方程不涉及人的直观感受。它涉及的是在实验室当中测出的特定量，如加速度、能量、力、质量和光速等。这些量不是从某处摘下来、挖出来或者订购来的，也不能像苹果或皮球那样放在手中。而且方程有一种诗歌所没有的结构：它可以表述一组量等于（在更广泛的意义下，也可以是大于或小于）另一组量。

方程所涉及的这些量并不是总可以分辨清楚的。有个故事说的是陆军上尉要选出一位中尉，向三位候选人提出了同样一个问题：“1 加 1 等于多少？”

第一个人回答说：“自然是 2 了。”

第二个人回答说：“这得看 1 代表的是什么了。1 可能代表向量，这样

它的值就可以在 0 到+2 之间任意选取。”

第三个人回答说：“你希望结果是多少？”

很显然，读者可能已经猜到了，结果是第三个人被选作中尉。在这个故事中，第二个人也有点意思，他采用了多义的概念，就是说 1 既可以看作数字，也可以看作向量的模。但是，正如故事中所讲的那样，方程与世界之间的联系并非如看上去那般简单。方程指明了特定量之间的关系，指出新的事物，给人以力量，改进人们观察问题的思路，因而在诸多方面改变了人们对世界的看法。

[8] 弗兰克・维尔泽克，“Whence the Force of F=ma? I:Culture Shock”，《今日物理》（*Physics Today*），2004 年 10 月，第 11～12 页；“Whence the Force of F=ma? II:Rationalizations”，《今日物理》，2004 年 12 月，第 10～11 页；“Whence the Force of F = ma? III: Cultural Diversity”，《今日物理》，2005 年 7 月，第 10～11 页。

目　　录



文明的基础
毕达哥拉斯定理

$$c^2 = a^2 + b^2$$

说明：直角三角形斜边长度的平方等于另两边长度的平方和。
发现者：不详。
发现时间：不详。

直到今天，毕达哥拉斯定理仍旧是整个数学中最重要的一个。

——J. 布罗诺斯基（J. Bronowski），
《文明的跃升》(*The Ascent of Man*)

毕达哥拉斯定理的发现过程已经无法考证了，不过关于该定理的二次发现却有数不清的故事。故事的主角有些是该定理的教授者，有些是发现该定理的人。毫不夸张地说，这些故事有时候改变了这些发现者的命运和职业生涯。毕达哥拉斯定理的魔力在于：虽然定理本身比较复杂，并不显而易见；但证明的过程却是高度浓缩的，也是一种独特的体验。

受毕达哥拉斯定理影响而改变人生轨迹的人当中，有一位著名的政治哲学家，叫做托马斯·霍布斯（Thomas Hobbes）（1588—1679 年）。40 岁之前的霍布斯虽有天分，却没有什么原创性的思想。他擅长人文，

却以为自己算不上博学。他的主要成就是漂亮地翻译了古希腊历史学家修昔底德的著作。译作里虽不时会出现讹误，不过都还无伤大雅。虽然在霍布斯的时代，开普勒、伽利略和其他一些人已经作出了一些激动人心的重大突破，引爆了学术界的革命，不过霍布斯本人却基本没怎么接触科学。

一天，霍布斯在路过一位熟人的藏书室时，发现桌子上有一本摊开的欧几里得的《几何原本》（*Elements*）。这在当时不足为奇。那时候的绅士，如果能收藏一卷体面昂贵的重要作品（如圣经），往往不会将其束之高阁，而是放在外边让来访者随意浏览。通常他们还会将书摊开放，以展示其中著名的章节，如圣经中的赞美诗。

欧几里得的《几何原本》就是一本圣经。它以定理和假设的方式展示了当时已有的大部分数学知识。这本书自公元前 300 年问世以来，学者们就一直没有中断过对它的分析和研究，书中的内容在霍布斯的时代仍旧适用。那时，《几何原本》是除圣经之外，流传最广、人们研究最多的书。霍布斯所看到的章节正是第 1 册的命题 47，毕达哥拉斯定理。

命题的内容是：直角三角形斜边长度的平方等于另外两边长度的平方和。霍布斯看到这个命题时非常惊讶，甚至说出了亵渎神灵的话来。以至于他的朋友、第一个为他写传记的作家约翰·奥布里（John Aubrey）没有把霍布斯当时所说的话全写出来——“老天爷（作证），”霍布斯发誓说，“这不可能！”[1]

霍布斯来了兴致，继续读了下去。循着命题 47，他又去看了书中的其他命题：命题 46、14、4 和 41。而这些命题又引用了其他命题。霍布斯一一读完，最后很快就确信，起初看似惊人的定理是正确的[2]。

奥布里写道：“从此霍布斯就爱上了几何学。”这表明霍布斯的人生轨迹也随之发生了改变，他迷上了绘制几何图形，在床单甚至自己的大腿上作演算。他全身心地投入到数学中，并显示出一定的天分。不过他的数学能力还只能算是中等。霍布斯还卷入了数学纷争，他在那些无望

的数学圣战中所表现出来的态度，至今还让他的一些传记作者和“粉丝”觉得难堪[3]。这些事情并没什么意思。重要的是一个定理就使霍布斯和他自己的学识发生了转变。有一位批评家这样描写霍布斯与毕达哥拉斯定理的第一次邂逅：“他以后的思想都因毕达哥拉斯定理而发生了改变。”[4]

于是，霍布斯开始批评当时的道德和政治哲学家，认为他们缺乏严密的思维，受前人的影响太深。他还不合时宜地将这些哲学家与数学家进行比较，认为数学家工作虽然做得慢一些，却是从人人都明白和接受的“最低原理”出发的。在《利维坦》（*Leviathan*）等书中，霍布斯开始用类似的方式重建了政治哲学。他先是清晰、准确地给出术语的定义，之后依次推导出它们的深层含义。霍布斯从毕达哥拉斯定理中学到了新的推理方法，以及如何令人信服地呈现由推理得出的结论。这些推理方式既是必要的，也是通用的。

毕达哥拉斯定理：法则

人们普遍用“毕达哥拉斯定理”一词来指代两种情形：一个是法则，另一个是证明。法则叙述的仅仅是事实，说明直角三角形各边长之间的等式关系：斜边长度的平方（c^2）等于另两边长度的平方和（a^2+b^2）。该法则有实际的价值，例如，如果人们已知三角形两直角边的长度，就能利用该法则计算出斜边的长度。不过证明就不同了，它陈述的是人们如何知道某个事实是正确的。

这种双重涵义所带来的意义上的混乱，全是因为“定理”一词。该词可以表示已经（或人们认为能够）被证明的结论。它来自希腊语，意思是“看”或者“考虑”，与“剧院”（theater）一词的词根相同。霍布斯等人第一眼看到毕达哥拉斯定理之后，注意到两件截然不同的事情：一个是结果、法则或者被证明的事物——斜边定律；另一个就是证明过程，即人们了解证明的途径。

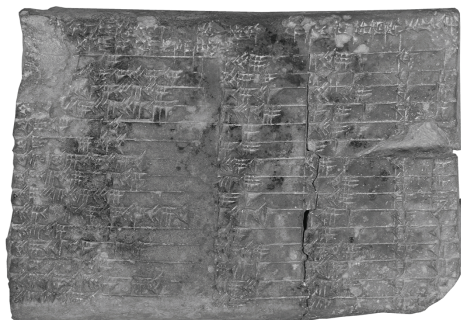
毕达哥拉斯定理极其重要，它对描述空间起着至关重要的作用。木匠、建筑师和测绘师在建设小型和大型工程时都离不开它。石匠会（据说是源自中世纪石匠工会的秘密组织）将毕达哥拉斯定理作为标志，也是因为这个原因。有篇石匠资料引用了毕达哥拉斯定理，认为它“包括或代表了砖石建筑和文化的基础”。在共济会的地毯上常常绣着简化版的欧几里得图形证明（称为“经典形式”）。它还适用于天体空间，因而对航海学和天文学非常重要。

早在欧几里得之前，甚至毕达哥拉斯之前，毕达哥拉斯定理就已经存在了。古代的工匠就通过经验发现，直角三角形的三条边的长度是由特定的数组构成的，例如 3、4 和 5 或者 6、8 和 10。这些三元数组称为“毕达哥拉斯三元数组”。它们虽然简单，却有着重要的实际意义。因此不同国家地区的人们各自独立地得出这一发现也就不足为奇了。古代的另一个发现就是这些三元数组之间存在着 $c^2=a^2+b^2$ 的关系。公元前 1800 年，巴比伦人的楔形文字泥板上就已经有了由 15 行毕达哥拉斯三元数组组成的表。这块泥板现收藏在哥伦比亚大学，称为普林顿 322 号（Plimpton 322）。毫无疑问，该表是三角学的表，也可以说是按照法则计算直角所对斜边的教学辅助工具。表中不包括变量，不过似乎是在通过一系列的例子对定理进行验证[6]。

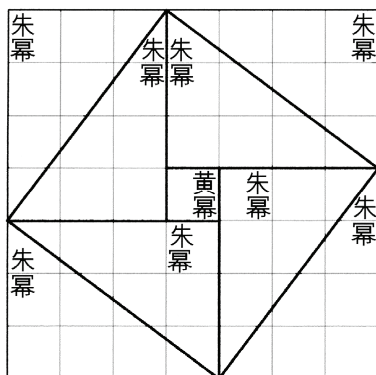
古印度人也已经知道这个定理。从《绳法经》（*Sulbasūtras*）一书中可以找到它的应用。该书与佛经同时问世，约成于公元前 500 年到公元前 100 年，不过它所传授的知识却是更早以前的了。虽然该书的表达常常是不正式的、粗糙的，也没有提供什么证明，但它却为宗教建筑的建造提供了相当多的几何知识[7]。

中国现存最早的有关天文学和数学著作是《周髀算经》。书中的文字可以追溯到公元前 1 世纪，而书中的内容据说还要早几个世纪。该书也包含了毕达哥拉斯定理的内容。这一定理的一个应用实例是计算太阳与地球之间的距离。推理的过程涉及了竹竿和竹竿的影子，并假定地球是

平的。历史学家认为《周髀算经》之所以著名，是因为它“第一次从理性的角度，完全以数学的形式解释了地平说”。[8]现存最早版本的《周髀算经》中有一张常以棋盘为背景出现的图。从该图上可以清楚地看到以斜边为边长的正方形面积等于以另两边为边长的正方形面积之和。不过，几乎可以确定，这幅图源自公元 3 世纪，远在欧几里得之后。



公元前 1800 年，巴比伦人的楔形文字泥板上就已经有了由 15 行毕达哥拉斯三元组组成的表。这块泥板现收藏于哥伦比亚大学，称为普林顿 322 号（Plimpton 322）。毫无疑问，该表是三角学的表，也可以说是按照法则计算直角所对斜边的教学辅助工具



《周髀算经》最新版本中的图，汉字指的是正方形的颜色

毕达哥拉斯定理作为数学知识一部分，在巴比伦人的泥板、印度的《绳法经》和中国的《周髀算经》中，都是以实际应用的形式展现出来的：普林顿 322 号是出于教育的考虑，《绳法经》是出于宗教的考虑，《周髀算经》则是出于天文研究的考虑。古书都没有给出该定理的直接明确的证明，而只是把它作为一种计算距离、验证结果的方法。不过形式偶尔还是有严谨的时候。

毕达哥拉斯定理在数学上众多的里程碑中无疑是独一无二的。它的表现形式丰富多彩，平淡中包含诗意。在几千年的历史长河中，它被用于土地丈量、运河开挖、晾衣绳搭设、人行道铺设、马路和沟渠建造等诸多方面。有一篇埃及的手稿这样写道："塔上倚着一个梯子，梯子长 10 腕尺[①]，底部离墙 6 腕尺，问梯子有多高。"还有一篇中世纪时的手稿是这么写的："墙上倚着一个长 20 英尺（1 英尺=30.48 厘米）的矛，如果矛的底部向外移动 12 英尺，那么此时矛在墙上所倚的高度是多少？"印度的书是要读者计算池塘的深度，池塘上还有红鹅游来游去。问题是这样的：荷花苞尖一开始在水面之上 9 英寸（1 英寸=2.54 厘米）的位置，突然刮来一阵风，荷花苞尖被吹到了水下 40 英寸的位置，问池塘的深度是多少？当然，荷花的茎是扎在水底的。这样的问题使数学变得有趣。

毕达哥拉斯定理已经成为人类知识的典范，对它的了解体现了人类的智慧。在电影《绿野仙踪》（*Wizard of Oz*）的结尾，为了证明自己的确有头脑了，稻草人（Scarecrow）笨手笨脚地将这条定理加以改编："等腰三角形任意两边的平方根之和等于另一边的平方根。"这种轻率的陈述真是绝了，我们作为观众自然也不会去学。就让它一直留在童话里吧。

毕达哥拉斯定理：证明

然而，定理的证明和定理本身有着天壤之别。证明需要从第一原理出发，给出结果的一般有效性。它研究的是定理本身，与定理的实际用途

① 古代的长度单位。——译者注

无关。它侧重于结论推导的过程（这样才能令人信服），而非结果。证明所叙述的是人们理解方程的过程。因此，要给出定理的证明，就要以不同视角看待数学，而不能简单地陈述定律。证明不是对权威的维护，而是对真知的认可。证明不是简单地把前人的智慧作为代表作传授下来。恰恰相反，它应该是天才之笔。证明并不是说“事实就是这样”或者“天才告诉我们就是这样的”。相反，对结果的证明是任何人都可以参与的“旅程”，至少从原则上来说是这样。这得益于人们目前已有的诸多数学定义和数学概念。因此，实际上定理的证明是说：“照着这个做，就能发现其实我们已经知道了要推出结果的全部步骤！”因此，定理的证明实际上是建立了一个路标，任何人只要循着这个路标指示的路线走下去，就能到达终点。人们可以对此有充分的信心，指导自己踏上更多探索未知的道路。正是有了通过证明关键方程而建立起的路标，数学才从复杂的地貌变成了一幅风景。数学的其他部分仍旧存在，只不过它们是在风景的背景中而已。

虽然传统上，人们认为斜边定律的第一个证明是毕达哥拉斯（约公元前 569—前 475 年）给出的，不过这个说法是在其后 500 年首次提出的。其实事实并非如此[9]。该定理的证明思想起源于古希腊时代，经历了几百年才形成的。这一阶段的顶峰是欧几里得的《几何原本》。《几何原本》完全以明确、正式的证明形式把数学呈现出来。第 1 册的倒数第 2 个命题就是毕达哥拉斯定理的证明：直角三角形中，直角对边边长的平方等于另两边的平方和。第 1 册的最后一个证明（命题 48）则是其逆定理：如果三角形某条边边长的平方等于另外两边的平方和，则该三角形是直角三角形。证明如下：沿着三角形的三条边分别画出三个正方形。从直角的顶点出发，垂直于斜边画一条线，并延长到斜边所对应正方形的对边。这样，将大正方形分割成两个矩形，每个矩形的面积就分别等于两个小正方形的面积：两个小正方形的面积之和就等于斜边上的正方形的面积。有趣的是，欧几里得的证明与图中直线构成的图形联想到一起。人们按图形暗示出的有趣形状，把它称为风车证明法、孔雀证明法或花轿证明法。

说明欧几里得《几何原本》中一个证明的经典图形

任何一个伟大发现的背后，似乎都有一种难以抑制的冲动：去看看之前是否也有人提出了该发现，或者虽然发现了却没记录下来，或者是与发现擦肩而过。毕达哥拉斯定理（似乎我们注定要这么叫了）也不例外。历史学家发现，人们证明毕达哥拉斯定理的能力似乎是文明进步程度的一种标志。他们根据普林顿 322、《绳法经》、《周髀算经》和其他资料，对巴比伦、印度和中国发现毕达哥拉斯定理的情况进行研究[10]。不过在研究过程中，很容易混淆或忽视毕达哥拉斯定理的经验法则和毕达哥拉斯定理的证明两者的不同。

新证法

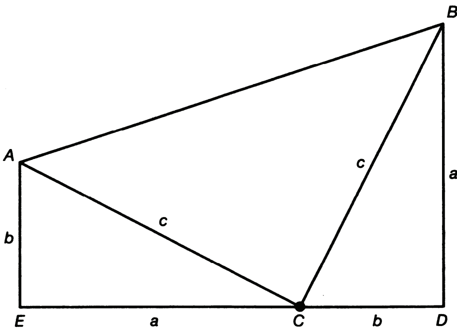
人们有时候会自己踏上"旅程"，去发现毕达哥拉斯定理，而并不依靠教师的帮助。法国数学家和哲学家布莱斯·帕斯卡（Blaise Pascal）就是这样一个人。帕斯卡的父亲不允许他在家中讨论数学，怕会影响他的希腊语和拉丁语学习，认为语言的学习最重要。但小帕斯卡凭着一根炭笔就开始了几何学的研究。他在欧几里得《几何原本》中找到了许多证明，其中就有毕达哥拉斯定理[11]。

发现定理的新证法是完全可能的。如果说在数学的诸多里程碑中，毕达哥拉斯定理所囊括的应用和实例范围是独一无二的,那么它的证明法之多也同样是独一无二的。大多数证明虽以相同的公理作为基础，但证明的过程却各有不同。这中间的许多证法，特别是苏格拉底、欧几里得等给出的最早的证明，以及之后《周髀算经》中给出的证明采用的都是几何方法。其中，a、b 和 c 指的是三角形各边的边长。证明的方法是通过几何运算，得出面积之间的关系。其他证明采用的是代数法，以复杂的数学（数字代表抽象的事物，甚至可能是向量）为基础。有些所谓的证明，假定结果可以通过毕达哥拉斯定理得到证明，最终变成了循环论证。代数方法（巴比伦人理解该方法）产生了毕达哥拉斯定理的 $c^2=a^2+b^2$ 的形式。

公元 4 世纪，希腊（亚历山大里亚的）几何学家帕普斯发现了一个定理，扩展了欧几里得几何。几个世纪之后，住在巴格达的阿拉伯数学家泰比特・伊本・奎拉（Thabit Ibn Qurra，836—901 年）提出了几个新的证法，并修改了先前的《几何原本》译本。两个半世纪之后，印度数学家婆什迦罗（出生于 1114 年）为《周髀算经》中证明方法的简洁所深深吸引。他随后采用简单的图形形式，对《几何原本》进行了修改。新版《几何原本》的证明里不再有文字解释，只有两个字“见图”。

后来，意大利艺术家列奥纳多・达芬奇（Leonard da Vinci）、荷兰科学家克里斯蒂安・惠更斯（Christiaan Huygens）和德国哲学家特弗里德・莱布尼茨（Gottfried Leibniz，1646—1716 年）都给出了新的证明。1876 年，美国国会议员詹姆斯・加菲尔德（James Garfield）也给出了证明。他后来成为美国第 20 任总统。有关毕达哥拉斯定理证明的集子已有十多种：1779 年，在巴黎出版了 18 种证明法；1880 年，德国出现了提供 46 种证明法的专著；1914 年，在荷兰出版了 96 种证明法。《美国数学月刊》（*American Mathematical Monthly*）是科学家一般都会关注的杂志，自 1894 年创办之日起第一期就开始刊登毕达哥拉斯定理的证明。该杂志略带些傲慢地提到：问题求解“是数学研究的最低级形式之一”，用于应

用，却没有科学价值。但该杂志却“专门开辟版面用于问题求解”，诸如毕达哥拉斯定理之类，起到教育的作用。“问题求解是引导思维进入更高级的原创性研究领域的阶梯。许多原本智力平平的人在掌握了某一个问题求解后，跨入到研究的行列中。”[12]1901 年，在发表了毕达哥拉斯定理的第 100 个证明之后，《美国数学月刊》的编辑放弃了，宣布“该定理的证法是无穷无尽的，本刊今后将不再接受此类稿件”。



詹姆斯·加菲尔德给出的毕达哥拉斯定理证法的示意图

但有一个人仍乐此不疲。这个人叫伊莱沙·卢米斯（Elisha S. Loomis），职业是教师。他是《美国数学月刊》的订户，也曾给出过一些证明。卢米斯一直都在收集定理的证法。这些证法很多都是由一些颇为聪颖的年轻人的老师转交给他的，这些教师知道卢米斯对此颇有兴趣。1927 年，卢米斯（此时的他已是大学教授）出版了《毕达哥拉斯命题》一书，收集了 230 个证法；1940 年，87 岁的卢米斯又出版了该书的第二版，收集了 370 个证法[13]。他将两本书都献给了共济会分会。他依据证明的内容，把它们分为几何证法、代数证法、动态证法和四元数证法。其中的大部分都是几何证法：第 31 种证法是惠更斯给出的，第 33 种是欧几里得给出的，第 46 种是达芬奇给出的，第 225 种是婆什迦罗给出的，第 231 种是加菲尔德给出的，第 243 种是《周髀算经》给出的。在代数

证法中，第 53 种是莱布尼茨给出的。卢米斯对迷上了证明过程的学生们敢于挑战，并尝试想出新证法的做法很是欣赏。他喜欢去发现有趣的人给出的有趣证明，或者对年轻的证法提供者加以奖励[14]。卢米斯认为有些人不尊重该学科，不同意这些人的看法。美国的几何学教科书因为略去了欧几里得的证法，也遭到卢米斯的嘲笑。他认为欧几里得的证法所体现的可能是“原创性或独立性”。他还挖苦说：“（几何学教科书中）没有欧几里得的证法就好比是哈姆雷特剧中没出现哈姆雷特的影子。”[15] 他第二版的书中的最后一句话是“新的证法没有尽头”[16]。

卢米斯是对的，新的证法的确层出不穷。在《吉尼斯世界纪录大全》(*Guinness Book of World Records*)网站上的“毕达哥拉斯定理的最新证法”一栏中显示的是一位发现了 520 种证法的希腊人。在此时，不知又有多少新的证法已经出现了。

证明的奥妙是什么

所有这些证法都引发了两个问题。第一个问题是：一种证法难道还不够吗？我们都知道一种用途是远远不够的—— 一个法则，最关键之处在于它能适用于各种不同的情形。但是证明呢？只有毕达哥拉斯定理的几个证明是对欧几里得已经证过的问题进行概括和充实。然而，卢米斯所收集的证法中的绝大部分都不是这种类型。而且这些证法也没有使这些已经很明确的结论变得更明确。但除了发现本身之外，证明的迷人之处还在于观察发现可能采用的视角。这样一来，那些隐藏的可能性和假定的结果就可以成为事实。科学的目的是使世界变得更加丰富，拓宽其外在表现形式，阐述世间事物的本质。随着科学的进步，世界的范围也在拓宽。

第二个问题是：为什么人们会把注意力集中到这样一个特殊的定理上？为什么几千年来它令业余爱好者和专家学者如此着迷？一部分原因归于个人的经历：毕达哥拉斯定理是大多数人第一次接触到的深刻证明，其结论在证明之前并不是显而易见的，正如霍布斯当时的情形。这是人

们的第一次数学发现之旅，从另一头出发发现了崭新的事物。很早的时候，人们也发现了其他的一些漂亮的证明，例如 2 的平方根是无理数，素数有无限多个等，所以这只能是答案的一小部分。人们也知道一些与毕达哥拉斯定理类似，却更加高效、更加有用的定理（如欧几里得《几何原本》第 6 册中的命题 31）。不过这些定义远没有毕达哥拉斯定理那般引人注目。一个明显的例子就是余弦定理：$c^2=a^2+b^2-2ab\cos\theta$。该定理适用于包括直角三角形在内的所有三角形，涉及三角形的两条边和一个角的余弦。毕达哥拉斯定理只是余弦定理的特例。然而，余弦定理没有什么迷人之处，部分原因可能是要证明它就得熟悉三角，所以很难想象霍布斯那样的人会因为这样一个定理就发生什么转变。

要完整回答毕达哥拉斯定理的魅力得从三方面着手：定理的应用场合是显而易见的；定理的证明是可行的；证明定理的过程提升了人们思考问题的高度，让人们体会到了思考的乐趣。

首先，毕达哥拉斯定理刻画了空间。它不仅适用于木工手艺、建筑学、物理学和天文学，也适用于几乎所有的职业领域和实际应用场合。在三维空间中，用毕达哥拉斯定理表示的距离表达式是 $\sqrt{x^2+y^2+z^2}$，例如鞋盒的对角线；在四维的欧几里得空间中，距离的表达式是 $\sqrt{x^2+y^2+z^2+w^2}$；采用爱因斯坦广义相对论的闵可夫斯基解释，四维时空版本的距离表达式是 $\sqrt{x^2+y^2+z^2+(ct)^2}$，其中 c 是光速。对上式进行适当变形，就能得出描述物体中分子三维运动情况的热力学方程。毕达哥拉斯定理还可用于广义和狭义相对论。狭义相对论用它描述观察者在某一参考系中观察另一参考系中光线传播的路径。广义相对论复杂一些，人们用它描述光线在弯曲的四维时空面上传播的情形。该定理在更高深的数学中得到进一步推广。伊莱·马奥尔（Eli Maor）在其所著的 *The Pythagorean Theorem: A 4,000-Year History*[1] 一书中称毕达哥拉斯定理

① 本书中译文名为《勾股定理：悠悠 4000 年的故事》，已由人民邮电出版社出版。
——编者注

是“数学中最常用的定理”。[17]这不仅是因为该定理有着直接的应用，也与马奥尔把它称为“毕达哥拉斯定理幽灵”有关。其他的许多表达式都是直接或间接由毕达哥拉斯定理推出来的。一个例子就是著名的“费马大定理”，最终于 1994 年证明。其内容是对大于 2 的任何整数 n，方程 $a^n=b^n+c^n$（所有变量都是正整数）均不能成立。该定理是否定性的，没有断言恒等关系的成立，所以无法用方程形式表示。

其次，霍布斯的经历表明，即使毕达哥拉斯定理包含了一些在证明伊始看似难以置信的数学知识，人们也可以在没有接受过任何数学训练的情况下，用简单而又令人信服的方式加以证明。这也正是自柏拉图以来的哲学家和科学家将其作为推理典范的原因所在。在《世界体系》（*On the World Systems*）一书中，伽利略援引毕达哥拉斯的证法，解释了事实与证明之间的差别（现在称为发现与证实之间的差别）[18]。在《指导心灵的规则》（*The Rules for the Direction of the Mind*）一书中，法国哲学家、科学家雷内·笛卡儿（Rene Descartes）援引毕达哥拉斯定理，说明了采用符号和记号的好处，并把它们引入到了数学中。黑格尔认为“证明高于一切”。证明阐述了几何学的意义，使它以科学的方式继续发展。对黑格尔而言，就是如何表明相同之中有不同[19]。德国哲学家亚瑟·叔本华（Arthur Schopenhauer）是欧几里得证法的批评者之一。叔本华认为该证明具有典型性是有另外的原因的。他嘲笑这种证法就好像是“捕鼠器”，先是诱导读者，之后再给他们“下套”。他认为证明的逻辑性虽然没有问题，但是过于复杂，显得不够直观（叔本华本人喜欢直观的证明）。他认为欧几里得的证明就是一个典型的误导读者的例子。实际上，叔本华认为欧几里得给出的毕达哥拉斯定理的证法是违背当时哲学观的一个典型，它过于强调纯粹的逻辑，把逻辑置于洞察力和后天直觉之上。在叔本华看来，黑格尔的哲学体系在概念上不过是一个巨大的捕鼠器[20]。

再次，毕达哥拉斯定理让人们了解了发现的全部。人们在证明该定

理的时候，可以说并没有“学到”什么东西，因为我们从小就知道毕达哥拉斯定理。但是当证明一步步地进行下去时（当人们把问题放在一个大背景下，当细枝末节相互因为必然整合到一起时），人们似乎离开了原地，去了另外一个地方。那里是一个比现在要古老得多的真理王国，不管人们身在何方，只要付出一点努力就能到达。在那里，直角三角形不再是特殊的，所有事物都是相同的，人们无需从头证明就能了解到这一点。三角形只是一个特例，在它的后面隐藏着别的事物。这种体验是愉悦的，甚至是令人兴奋和难忘的。证明是以先前未有的语言形式给出的问题答案。令人不可思议的是，这种语言在你感觉已掌握它的那一瞬间便出现了。如果没有那一刻的洞察力，毕达哥拉斯定理就仍然是代代相传的权威，而不能成为一个有深刻见解的证明。

《美诺篇》（*Meno*）中的毕达哥拉斯定理

毕达哥拉斯定理有如此之魅力的原因有三点。柏拉图在对话《美诺篇》中已经对这三点进行了明确的阐述。该书成于约公元前 385 年（毕达哥拉斯之后约 100 多年，几乎早于欧几里得的《几何原本》100 年），是已知的最早、最有名和最复杂的叙述毕达哥拉斯定理证法的书。它第一次广泛地阐述了古希腊现存的数学知识。书中叙述了苏格拉底如何哄骗一个不懂数学的奴隶小孩证明毕达哥拉斯定理的一个特殊情形——等腰直角三角形。

主要参与者是苏格拉底和美诺。美诺是一个英俊的塞萨利青年。他性格急躁，拒绝思考难以理解的想法，总是希望答案能令人印象深刻。对老师来说，这简直就是噩梦。在如何学习美德的问题上，美诺让苏格拉底极为苦恼。苏格拉底发现想让美诺思考是件很难的事，真是人如其名，他的名字的意思就是“不让步”或者“不回头”。“教育”一词字面上的意思是“出发”。苏格拉底发现自己没法引领美诺。

有一次，美诺认输了，并问苏格拉底，怎么可能把所有的知识都学到手。这就是著名的美诺悖论。如果你不知道自己要学的是什么，那么即便你遇见它也还是会与它擦肩而过。但是，如果你已经知道了，那何需费力去寻找呢？美诺意指任何尝试都是没有结果的。

正如今天的哲学家所说的，美诺悖论之所以会产生，是因为他错误地假定知识是相互分离的。实际上，人们正是因为有了现在的知识，才会去注意那些未知的知识。用已知去发现未知，就能拓宽知识的范围，使知识体系更加充实连贯，消除知识间的隔阂。但该过程会不可避免地引入一些新的漏洞和缺陷。获取知识的过程并不是把别人给自己的东西都一股脑地扔到思想库中去。它是一种在部分和整体之间不断反复的过程，基于已知发现新事物，扩大人们赖以理解世界的基础。

当然，苏格拉底教育美诺采用的并不是这种方式。那么深奥的道理是美诺不能理解的。他深知年轻人易于上当的特点，所以采用了一种哄骗的方式。苏格拉底是这么说的："我告诉你一个圣人都信奉的传说吧。他们说灵魂是不朽的，所以看见并了解了世间的万物。因此，其实人们的内心深处早已了解了所有的一切，但是因为在尘世间停留的时间短暂，把它们全忘却了。如果人们的精力允许，就应该克服这种无知，重新了解一切。"

苏格拉底通过这样一个传说，以一种诗意的方式告诉美诺，学习既不是被动接受其他人传授给你的知识，也不是不假思索就去相信法则。学习是精力集中的主动过程，需要激励自己去发现，始终保持主动的思维。当你发现某个事物是对的时，就会发现它似乎早已在那里，是知识体系中的一部分，只是自己没有注意到罢了。它深深地印刻在你的头脑中，就好像你早就知道了一样。为了学习知识所作的准备、练习和证明只是让你想起知识。这正是比喻的真谛。

美诺很喜欢这个传说。但他还没有把握住关键，请苏格拉底加以解释。苏格拉底用了一种新的办法为美诺进行实际演示。他让美诺叫来一

个奴隶，“随便叫一个就行”，美诺答应了。之后，苏格拉底就哄骗这个没有受过任何数学训练的奴隶小孩来证明这样一个几何定理：以某个正方形的对角线为边构成的正方形的面积，是该正方形面积的两倍。这其实是毕达哥拉斯定理在等腰直角三角形情况下的形式。苏格拉底在沙子上一步步画出图形，让美诺仔细听好并进行监督，防止他加进任何数学信息，以保证这个小孩最终得到的结论“只是被他自己回想起为的”，而不是被别人灌输进来的。

现在的读者会认为后面所发生的肯定是骗局。他们会认为苏格拉底在幕后操纵，愚弄奴隶小孩，让他跟着做口形对词。现代人觉得把学习看作回忆是很荒谬的。他们认为，真正的学习是把新的信息“下载”到大脑中，之后辅以作业练习进行强化。但是如果仔细去读柏拉图的书，就会发现奴隶小孩的学习才是真正的学习。他学习的所有内容都能回归到最基本的知识。苏格拉底保证所有新的认识都是基于奴隶小孩已有的经验。在这里，我们看到的是奴隶小孩正走在学习毕达哥拉斯定理的小小旅途中。可行的途径有千万条，应该走哪一条呢？苏格拉底告诉他该走哪条路，并给他选择道路的动力。

“知道什么是正方形吗？”苏格拉底在沙上边画边问，“是像这个，四条边都相等的图形吗？”奴隶小孩说：“是的。”

苏格拉底又问：“那你知道怎么把正方形的面积变成之前的两倍大小吗？”

“当然知道了，”奴隶小孩回答，“把边长都变成原来的两倍就行了。很显然的事情！”

这自然是错的，不过苏格拉底仍然不露声色。一个好的老师应该让学生自己看到错误的所在。当奴隶小孩把正方形所有的边都变成了原来的两倍后，马上就发现自己错在哪里。新的大正方形的面积是原来的四倍，而不是两倍。

而苏格拉底只是说：“再试一试。”奴隶小孩就把边长变成原来的 1.5

倍。苏格拉底画出图形之后，小孩马上发现自己又错了。

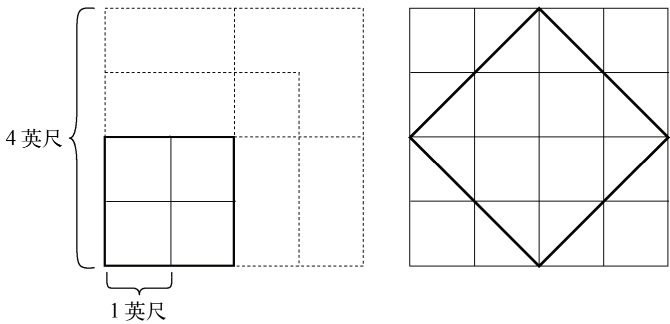
苏格拉底于是问小孩知不知道怎么才能把正方形的面积变成原来的两倍。很有意思的是，苏格拉底这么做其实是为了美诺。小孩说："我不知道。"

关键的时刻来了。第一，苏格拉底让奴隶小孩第一次发现了自己知识的有限，让他知道了还有自己不知道的事情。第二，他摧毁了小孩的自信心。小孩之前一定认为自己什么都知道，然而现在他明白了事实并非如此。当然，奴隶小孩也并非一无所知。他知道的并不少——问题的答案限定在一个很小的范围之内，比边长要大，比边长的 1.5 倍要小——这一点小孩虽然知道，可无法表达出来，这给他一种不好的感觉。答案已经有眉目了，可就是无法表达。这是因为奴隶小孩并不知道描述答案的语言。他起先认为自己有能力理解这样一个问题，可现在却发现事实并非如此，这令他很不舒服。这种困惑激起了好奇心，而好奇心是学习所必需的。他渴望自己能得到指引，准备踏上学习的旅程。他渴望去发现和思考。在方程诞生的过程中，读者将会反复看到这种渴望所起的作用——引导人们去探索新的事物。有时候引发这种渴望的只是一个偶然事件，比如落下的一个苹果、随口说出的一句话、难以理解的数据，或是两个理论间的不一致等。苏格拉底使小孩变得困惑，令他想要跟随自己（一种"引诱"，这在当时可以被看成一种罪行，足以立马让苏格拉底受到控告并被判处死刑）。

苏格拉底利用了小孩的困惑。他擦去了先前的图形，重新画了一个正方形，边长为 2 英尺。他又在正方形的旁边画了三个相同的正方形。之后苏格拉底在图形上加了一个新元素——连接两个对角的一条线，"学者称它为对角线"。对角线并不是全新的元素。奴隶小孩之前在地板的马赛克和墙面设计上已经看到过（这种体验已经告诉它接下来会发生什么），现在只是回忆一下。但在这里对角线却变成全新的了。它一下子就把问题置于更加广泛丰富的情境中，使问题的答案变得显而易见。由

此，问题得到了重新的表述。

苏格拉底“乘胜追击”，很容易地就让奴隶小孩发现以对角线为边构成的正方形的面积是原来正方形面积的两倍。



左图是苏格拉底首先画出来的，该正方形的边长为 2 英尺。苏格拉底问奴隶小孩如何使正方形的面积变成原来的两倍。奴隶小孩先是说把正方形的边长加倍，变成 4 英尺。不过这样一来，正方形的面积就变成了原来的 4 倍。然后他又说把正方形的边长变为之前的 1.5 倍（增加 1 英尺），可最终得到的正方形的面积还是太大，为 9 平方英尺。在右图中，苏格拉底引入了对角线。奴隶小孩此时认识到由对角线所围成的图形的面积恰好是原来正方形面积的 2 倍

他转向美诺，试着用另一种方式教育他。他问美诺：“奴隶小孩是不是一开始不知道，后来才知道的呢？”“是的。”美诺承认了。“我向奴隶小孩灌输什么了吗？”“没有。”“那么他是自己找到答案的吗？”“是的。”“这些新出现的思想虽然是新的，现在却梦幻般地成为了现实，”苏格拉底继续说，并提出更多的问题。这样的提问会让学到的知识牢靠而不会忘记。无论他身在何方，这些知识都是他宝贵的财产，并且他所掌握的知识是正确的，跟别人的没什么不同。（人们把这样的问题叫作“作业和练习”。）如果我们坚持认为美诺悖论是正确的，那么奴隶小孩要么知道，要么不知道。这样一来，即便他知道了也会遗忘，就像传说中说的那样。“是的。”美诺表示认可。苏格拉底说：“我不会断言传说中的一

切都是对的，但是我相信它里面一定包含一些真理。”

美诺现在也认可学习是可能做到的，于是话题转向了一开始提出的关于美德以及如何传授美德的问题。于是，苏格拉底和美诺开始讨论谁适合来传授美德。他们很快就发现找不出合适的人选。无论是品德良好的公民，还是备受尊敬的城市统治者都不合适。此时一位有钱有势，名叫安尼托的雅典人加入了他们的讨论。安尼托听说苏格拉底认为连品德良好的公民都不能理所当然地成为传授美德的老师之后，颇为生气。他警告苏格拉底不要“说人们的坏话”。实际上，若干年之后控告苏格拉底的人当中就有安尼托。苏格拉底被判处死刑。

在这场“戏中戏”中，我们能看到很多。你其实也是在一场旅途中。我们看到毕达哥拉斯定理在自己面前产生。我们看到了奴隶小孩学习毕达哥拉斯定理的过程。我们看到苏格拉底在引导小孩，但是小孩可被引导的前提是他自己有主动性。我们看到美诺也处在旅程中，看到小孩从无知到掌握知识。我们看到了求知的过程：遇到困难停滞下来的时候，可以添加新的元素，丰富原有的知识库。那条新的线（对角线）起先是没有的。但是它一旦出现，就变得和其他线一样明显，丰富了知识库，令解决问题的途径变得清晰起来。这是有史以来第一次系统准确地描述亲身体验教育的记录。

除此之外，柏拉图也告诉我们自身的处境。这出“戏中戏”向我们暗示，我们自己其实与奴隶小孩的情况类似，然而却未有幸得到苏格拉底的指点——没有人向我们提出正确的问题，帮助我们学到正确的新知识。在一定程度上，人类的处境会激发出隐含的问题，令自己产生不舒服的感觉。幸运的话，机会就会像对角线那样出现。但是，要对答案加以描述，往往需要一种之前不为人知的语言。这样一来，就得预先设计好这种更加精细的语言。我们得像小帕斯卡那样学会怎么去添加那条对角线。柏拉图还告诉我们，要不断地提问题。人们一般习惯于把学到的

知识固定下来，所以总是存在着真理变成虚幻，现实变成梦幻的危险。这就是苏格拉底公开抨击《斐德罗篇》（*Phaedrus*）的原因所在。他把《斐德罗篇》几卷书称为已经无法答话的“鲜活语言的遗孤”。唯一的办法就是不断提问，不断对经验提出质疑，并保持好奇心。

柏拉图还有最后一个法宝。他通过一件事情指出，人们在努力的过程中，常常会遇到两种危险。第一种危险就是因为学术懒惰造成的惯性，即现代版的美诺，他们认为人无法真正学到知识。人所能做的就是对已有的知识体系缝缝补补，哪怕这知识是新的，也不过是一种推断。第二种危险来自政治家及其党羽，即现代的安尼托。他们告诉我们爱国心和对统治者的忠诚重于科学上的发问。这两种人都是在否定人类的文化成就，只是方式不同。对第一种人，需要保持耐心；对第二种人，得小心谨慎，有时甚至要顺从他们。《美诺篇》是留传至今的情节最为复杂的一个短篇文学作品，柏拉图用学习毕达哥拉斯定理的故事告诉人们，通向真理的路途比人们通常所认为的要困难和危险得多。



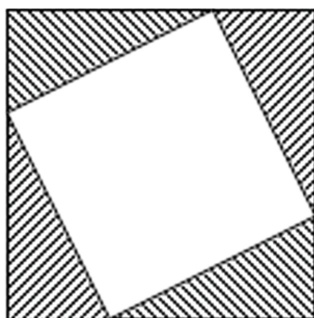
茶　歇

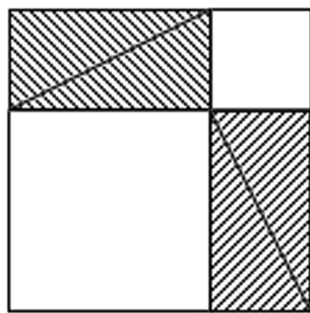
数学的法则、证明和魔力

人们都知道数学法则是怎么一回事，却不知道法则是如何证明出来的。拿毕达哥拉斯定理来说，它的证明方法有很多种，有的证法甚至一个字都没有。在欧洲最大的科学博物馆——法国的科学工业城（Cité des Sciences et de l’Industrie）里有一面立体的展示墙，上面就有毕达哥拉斯定理。上面有 3 个立体而中空的图形。每个图形分别紧邻着直角三角形的一条边。图形中装有一些有色液体。这些液体可以从一个图形流动到另一个图形中。展示墙转动时，液体就可以完全流到斜边上的图形中，

而其他两边的图形中一点也不会剩；反之亦然。欧几里得的《几何原本》19 世纪出版的版本，被人们认为是"本世纪最古怪，也是最漂亮的书之一"。这本书 1851 年曾在英国著名的水晶宫（Crystal Palace）中展出。现在在 eBay 上卖到几千美元。这本书采用彩线、彩图，对大部分证明中的文字加以浓缩提炼，几乎完全是用图形加以证明的，包括毕达哥拉斯定理的证明[21]。

哲学家大卫·索克（David Socher）想出了一个聪明的办法，来告诉学生毕达哥拉斯定理的法则和证明之间的不同[22]。他手上拿着 1 个大的白色正方形和 4 个有颜色的三角形，可他并不告诉学生自己接下来要做什么。"我只是说接下来要做个小小的演示。我要你们用不同的办法来移动正方形和三角形。这并不需要什么技巧，并不难，也不测试速度。仅仅就是一个很好玩的演示而已。"接着，他要求学生用两种方法把 4 个三角形（每个三角形的边长都分别为 3 英寸、4 英寸和 5 英寸）组合到正方形（边长为 7 英寸）上。学生们很快就能发现，两种方法中余下的空白面积是相等的。索克问这对于三角形有什么用呢？学生们说不上来。他接着问学生对三角形理解多少。这时不止一个学生提到毕达哥拉斯定理，可他们却没有发现眼前的几何图形与毕达哥拉斯定理之间存在的关联。索克接着写下了两个字："璞玉"。你看，再多解释一两句，法则和证明之间的关系就浮现出来了。





两种不同方法中的 4 个三角形

这样的时刻，即便是在成人看来，也是很难忘，甚至很渴望的。在

Quartered Safe Out Here: A Recollection of the War in Burma（《缅甸战争回忆录》）一书中，英国小说家乔治·麦克唐纳·弗雷泽讲述了二战期间向战友杜克演示毕达哥拉斯定理的故事。那个夜里，他们团正沿着通往仰光的公路挖掘掩体，他和杜克一边挖着，一边聊天。聊了香烟，聊了战争，聊了日本人，杜克谈兴正隆，提出让弗雷泽说点有文化的东西，也享受"一分钟不带脏话、有教养的讨论"。弗雷泽于是提出来"证明毕达哥拉斯定理"，杜克很开心，打趣说他觉得弗雷泽证明不了。（那天深夜，由于一系列的意外和误会，杜克被自己人当成了敌人，好几挺机关枪几乎把他打成了筛子，凄惨地倒在了前线的黑暗里。）

> 我用刺刀在战壕旁的地面上把我所知道的毕达哥拉斯定理的证法写给他看，这也是毕达哥拉斯本人的证法。期间出了几次错，有个地方忘了加垂线，不过最后还是对了，杜克脸上也露出了满意的表情。我"乘胜追击"，继续证明圆心角的度数等于圆周角的度数的两倍。杜克听得很用心，甚至有点超出我的想象。在四周都是日本人的漆黑夜晚，要如此全神贯注于三角形和圆形之类的东西实在不是一件易事[23]。

爱因斯坦曾在他的自传中写道：自己小时候根据三角形的相似性证明了毕达哥拉斯定理，欧几里得平面几何由此使他产生"好奇心"，并留下了"难以名状的印象"。"对任何一个第一次产生这种感觉的人来说，"爱因斯坦写道，"这都是一种神奇的经历——人们有能力通过纯粹的思考达到此种程度的确定性和纯粹性。"[24]

爱因斯坦的经历还表明毕达哥拉斯定理还能使人们产生另一种激动的心情。对那些不仅要知道如何证明，而且想要给出新证法的人来说，这种经历会告诉他们创造力是多么地激动人心。给出新证法的人可不像是看着剧情进展的观众。他们是剧作家，从事着数学家的工作，把数学当成一种创造性的艺术。他们体会着创作的乐趣，发现了数学的真谛就

是去作更多的数学。这种人已经发现了“发现的力量”。

对柏拉图、霍布斯、笛卡儿、黑格尔、叔本华、卢米斯、爱因斯坦、弗雷泽和其他数不清的人来说，毕达哥拉斯定理已经远远不只是一个计算斜边长度的方法。对那些惯于推理的人来说，会有一种超越纯粹结果的东西浮现出来。在经历某个事物时（比如量、数学），在某个时刻会有另一个东西——推理的思路，呈现在人们面前。它强健、不屈又固执，宗教之名赶不跑它，政治意识形态伪装不成它，学术技巧藏不住它。

1+1=2 告诉人们加法的概念，而毕达哥拉斯定理用类似的方式告诉人们如何作证明法。这样一来，哲学家们所谓的范畴直观（categorical intuition）就变成了可能：人们所看到的不再仅仅是内容本身，还包括思维框架。它所涉及的“旅程”是很短的。只要稍微瞥见旅途中的各个阶段，便可展现出知识的旅程。它是一个证明，论证了“证明”（Proof）的成立。

注　释

[1] John Aubrey，*Brief Lives*，编辑 Richard Barber（大不列颠：Boydell Press，1982 年），第 152 页。

[2] 这个故事看上去令人难以置信，不像是真的，尤其是在高呼“我发现了！”的时代。不过大多数传记作者还是相信的。人们在当时往往看不到有些事物的重要意义，只有过后才会意识到。最近一位为霍布斯写传记的作家 A.P. 马蒂尼奇（《霍布斯传》，第 84~85 页）对故事的真实性进行了有力的争辩。马蒂尼奇认为：“几何学对于霍布斯哲学的重要影响几乎不可能被夸大。影响霍布斯的并不是一大堆的几何公理、定理和证明，而是把一个事物与另一个事物用一种无可置疑的基础联系起来的方法。影响他的是几何方法，而不是几何内容。”

[3] 尽管他没有成为一个真正的专家，而且也会犯一些空有热情的业余爱好者经常犯下的错误。他曾经去研究一些不可能的问题，比如求出与圆的面积相等的正方形，对一个角进行三等分，以及得到一个体积是原来正方体两倍大小的正方体。霍布斯错误地认为自己成功了。

[4] Leo Strauss，《霍布斯的政治哲学：基础和起源》（芝加哥：芝加哥大学出版社，1959 年），第 29 页。

[5] Reid McInvale，“Circumambulation and Euclid’s 47th Proposi-tion”http:///www.io.com/~janebm/summa.html（于 2008 年 4 月 11 日访问）。另参见 James Anderson 的 *The Constitutions of Free Masons*（1723 年）：“伟大的毕达哥拉斯，被证明是欧几里得《几何原本》第 1 卷命题 47 的作者。如能对其妥善运用，可将其作为所有宗教建筑、民用建筑和军事建筑的基础”，（里士满，VA：Macoy，1977 年），第 203~204 页。

[6] O.N. Eugebauer 和 A. Sachs，“Mathematical Cuneiform Texts”，*American Oriental Series*，第 29 卷（纽黑文：美国东方学会，1945 年），第 38 页；Eleanor Robson，“Neither Sherlock Holmes nor Babylon: A Reassessment of Plimpton 322”，*Historia Mathematica* 28 (2001)，第 167~201 页。

[7] 包德哈亚那（Baudhayana，印度数学家）就说：“从矩形的对角线得出的面积等于两边分别得出的面积之和。”（David Smith 引用，*History of Mathematics*，第 1 卷（纽约：丹佛，1958 年），第 98 页。）不过他仅仅陈述了这样一个事实，并未对其加以证明。有一位学者写道：“我们必须要了解，他们感兴趣的只是有实际用途的几何事实。从而表述的形式就是最实际的了。”（G. Thibaut，《苏瓦经》（*Sulvasutras*）（加尔各答：Papatist Mission Press，1875 年），第 232 页。

[8] Christopher Cullen，*Astronomy and Mathematics of Ancient China: The Zhou Bi Suan Jing*（剑桥大学出版社，1996 年），第 xi 页。但正如 Cullen 所发现的那样，“整个过程更多的是文字而不是计算”，“虽然采用了欧几里得式的图标标记法，但在正文中却没有提到，反致误导”，不知作者所云。“书中基本没有称得上是计算的东西。”（第 80 页。）太阳的高度是 8 万里，即 4 万公里或 2.4 万英里（1 英里=1.609 公里）。之后中国又出现一本叫作《九章算术》（*Nine Chapters on the Mathematical Art*，约公元 250 年）的书，对法则进行了比较清晰的处理。《九章》所关注的依旧是实际问题，里面的第一章就是“土地丈量”。后续的章节主要是关于运河、税收等事务。最后一章是“勾股”（Kou ku），“勾”是“小腿”，指的是直角三角形的短边；“股”是“大腿”，指的是直角三角形的长边。“弦”就是连结两点构成的斜边。这一章包含了 24 个涉及直角三角形性质的问题。“但是证明并不是本书所关心的，”历史学家劳埃德（G.E.R Lloyd）说，“书中问题的数学推理方式更多的是基于类推和（问题、步骤和公式）结构上的共同点。”拿诗歌来说，

它最感兴趣的是关联、互补和对仗。这种风格在书中是没有的。

[9] 毕达哥拉斯时代的一个世纪之后，很多希腊作者就已非常熟悉斜边定律了。但这些作者都没有把毕达哥拉斯当成斜边定律的发现者。亚里士多德最擅于“论功行赏”，可他也没有提到毕达哥拉斯与斜边定律存在什么关联。证明的思想始于公元前 5 世纪。这一思想在公元前 4 世纪柏拉图讨论“说服”（persuasion）与“实证”（demonstration）之间的差异时达到顶峰。同一时期，亚里士多德讨论了证明的本质，欧几里得写出了《几何原本》。《几何原本》是第一本完全以证明形式呈现数学知识的书。早期的著作和之后希腊的正规证明思想之间存在着较大的差异。著作中数学知识的呈现，还是以实际结果作为出发点。但是劳埃德写道：“实际结果是一回事，证明又是另一回事。”“要给出定理或命题的正规证明，起码要保证步骤的准确和有效，然后才可能通过归纳和演绎得出定理或命题的正确性。”（劳埃德，《让智力不再神秘》（*Demystifying Mentalities*，第 73、74 页）劳埃德又继续写道，就目前已知，上述思想并非只由希腊由亚里士多德提出，在世界其他地方也出现了。比如有人就声称毕达哥拉斯定理在被毕达哥拉斯发现之前，就已在美索不达米亚、印度和中国被发现了。“从关键字句中我们看不出严格证明步骤和大概证明之间的差别。人们对两者的使用显然是不加区分的，说明作者并不关心结论的证明，只关注祭坛建造等实际问题。”（第 75 页）的确，生活在毕达哥拉斯（约公元前 569—公元前 475 年）之后 5 个世纪的作者认为毕达哥拉斯是给出斜边定律证明的第一人。但正如劳埃德指出的，这很可能是因为“后世的希腊注释者过于乐观地把功劳放到了英雄般的希腊哲学创建者身上”。（第 80 页）这些过于乐观的人中首当其冲的要数阿波洛道鲁斯。关于此人，人们知道得很少。只是知道他曾说毕达哥拉斯以牛祭祀这个“著名定理”的发现。许多作者于是听信阿波洛道鲁斯的说辞，如普卢塔克、阿特纳奥斯、第欧根尼・拉尔修、波菲利和维特鲁威。因为毕达哥拉斯学派对歃血仪式有着严格的限制，所以就有人对以牛祭祀表示了怀疑。于是普卢塔克等人又对故事进行了润色。“其实，”劳埃德推断（第 87 页），“对之后希腊科学发展最没有争议、最重要的是对知识框架进行系统说明的典范——欧几里得的《几何原本》所起到的作用。自此之后，几何证明被广泛采用。其范围除几何学外，还涉及光学、某些音乐理论、统计学和水力学、理论天文学的某些领域；它不仅可用于自然科学，也适用于生命科学的某些领域。”

[10] 举个例子。Otto Neugebauer 是破译了普林顿 322 中毕达哥拉斯三元数组的

第一人。他借用巴比伦人的泥板，认为泥板已经提供了“足够的证据，表明早在毕达哥拉斯 1000 多年之前，人们就已经知道‘毕达哥拉斯’定理了”。Otto Neugebauer，*The Exact Science in Antiquity*，普罗维登斯：布朗大学出版社，1993 年，第 36 页。

[11] Francs M. Cornford，*Before and After Socrates*（剑桥：剑桥大学出版社，1972 年），第 72~73 页。

[12] 《美国数学月刊》（*American Mathematical Monthly*），第 1 期（1894 年 1 月），第 1 页。

[13] 卢米斯，*The Pythagorean Proposition: Its Proofs Analyzed and Classified*。由 The Masters and Wardens Association of the 22nd Masonic District of the Most Worshipful Grand Lodge of Free and Accepted Masons of Ohio 出版，1927 年；*The Pythagorean Proposition: Its Demonstrations Analyzed and Classified*，安阿伯，密歇根：爱德华兄弟出版公司，1940 年。书的结尾是：“**最后的话**：毕达哥拉斯定理总是对的吗？黎曼和爱因斯坦分别于 1854 年和 1915 年对该定理加以推广，使其符合、包含欧几里得以外的几何学，并提出和验证了广义相对论。他们的努力似乎表明，毕达哥拉斯定理中蕴含的真理注定要成为一种基本因素。这种因素使过去、现在和将来的理论协调起来，共同构成宇宙的基本定律。”

[14] 第二版中的几何证明第 32 种“是 E. A. Coolidge 小姐给出的，她是一位盲人”；几何证明第 68 种“是第一个完全采用从给定三角形斜边出发的辅助线和三角形建立的证法。该证法由 Ann Condit 小姐提出并给出证明。她 16 岁，是印第安纳州南本德中心中学的学生，1938 年 10 月。这个只有 16 岁的女孩所做的工作，是连伟大的数学家，不管是印度的、希腊的，还是现代的都不曾作出的”；几何证明第 69 种“是原创性的。由西费城的中学生 Joseph Zelson 提出，经由他的叔叔送至本人手中。Joseph 能提出该证法说明他的智商很高”。卢米斯还说，这个证明和前面的一个证明“都证明了青少年也可以发挥演绎能力”；几何证明第 252~255 种“表明提出者具有高智商。还表明，不管是男孩还是女孩，只要通过独立的、逻辑的思考，就能做到这一点”；关于代数证法第 93 种，卢米斯评论说，该证法是由“19 岁的年轻人 Stanley Jashemski 提出的。他住在俄亥俄州的扬斯敦，智商极高”。

[15] 卢米斯，1927 年版，第 99 页。

[16] 卢米斯，1940 年版，第 269 页。

[17] 伊莱·马奥尔，《勾股定理：悠悠 4000 年的故事》，普林斯顿：普林斯顿大

学出版社，2007 年，第 xiv 页。

[18] 伽利略，*Galileo on the World Systems*，M. A. Finocchiaro 译，伯克利：加利福尼亚大学出版社，1997 年，第 97 页。

[19] G.W.F. Hegel，*Hegel's Philosophy of Nature*，第 1 卷，M. J. Perry 编辑、翻译（纽约：人文出版社，1970 年），第 228 页。

[20] 感谢我的同事 David Dilworth 的发现。

[21] Oliver Byrne 编，*The First Six Books of the Elements of Euclid*，伦敦：威廉-皮克林出版公司，1847 年。

[22] David Socher，“A Cardboard Pythagorean Teaching Aid”，*Teaching Philosophy*，28，2005，第 155~161 页。

[23] George MacDonald Fraser，*Quartered Safe Out Here: A Recollection of the War in Burma*，伦敦：哈泼柯林斯出版社，1992 年，第 150 页。

[24] *Albert Einstein: Philosopher-Scientist*，Paul Arthur Schilpp 编，伦敦：剑桥大学出版社，1970 年，见书中自传短文的开篇，第 9~11 页。

2

经典力学的灵魂 牛顿第二定律

$$F = ma$$

（牛顿的）说明：运动的变化与施加的力成正比，并且变化的方向沿着所施加力的方向。
发现者：牛顿。
发现时间：1684—1687 年。

牛顿第二定律 F=ma 是经典力学的灵魂。

——弗兰克·维尔泽克（Frank Wilczek），
《今日物理》（*Physics Today*）

方程 F=ma 是牛顿第二定律的简单表达式。这个方程就相当于经典力学中的 1+1=2。它看上去很直观，只是把平时的经验用定量的方式表达出来：对某物体施加一个力，物体要么开始运动，要么改变之前的运动状态。

是的，如果仔细研究 F=ma，就会发现它就像 1+1=2 一样不可思议。但实际上，F=ma 所描述的并非日常生活经验，而是没有任何阻力的抽象世界。在真实世界中，要保持桌子和车子以不变的速度运动，就得对它们持续施加推力。该方程不涉及著名的爱因斯坦质能互变关系，它所研究的中心内容是力。在相对论和量子力学等当代理论中，力的概念是

不存在的。最后，方程 $F=ma$ 既像是一个名字，也像是一种描述，看上去有点矛盾。它似乎是在定义力、质量和运动，又似乎是在描述人们所发现的这三者之间可测量的关系。

这样一个描述物体运动的简单方程怎么会隐藏着如此之多的复杂内容呢？读者如果仔细了解一下从古代一直到 17 世纪该方程的提出这一历史旅程，就不难找到答案。在这个方程出现之前，人们理解运动的方式有很多种，看待运动的视角也不尽相同。这些视角改变了人们对所观察到的现象的认识。在漫长的过程中，新的思想不断涌现，并占据主导地位，然后又从人们的视野中消失，如此更迭交替，直到某一天人们发现了全新的世界。

希腊关于运动和变化的观点

这个旅程早在人们认为世界受众神统治时就开始了。对理解事物而言，或许这是最简单、最直接的办法，因而不可避免，也很自然。人们从日常生活中常见的推力、拉力中分别获得了力的概念，例如通过肌肉的运动提起物体、挤压物体和使物体发生滚动等。早期的人把这种经验加以推广，就能理解大自然中的万物——从身边的雷雨现象到太阳恒星等遥远天体的运动等。他们认为这一切都是因为神灵施加了魔力。因此，他们的力的概念是与自然界中存在神灵这一宗教观念紧密联系在一起的[1]。

最早，人们希望能通过祭祀神灵，取悦他们，进而控制自然。这就是最早的技术。不过他们并没有得到预想的结果。要更有效地预测和影响自然，就得留意自然界中变化的类型和数量，比如四季的轮回、星体的各种运动、水火的特点等。但自然界中的变化实在是过于复杂多样。阳光白云、潮汐风浪、植物动物、男人女人、计划思想、房屋城市都处于永恒的生死存灭、涨落兴衰、颜色形式的变化和运动之中。作为人类，应该如何去理解这一切呢？

希腊哲学家亚里士多德（公元前 384—公元前 322 年）是已知最早对各种形式的运动和变化进行系统研究的人。他用同一个词“运动”（kinesis）来表示这二者。他认为运动很重要，理解了运动就相当于理解了自然界。他提出了一个包含各种运动形式的框架：有生命的和没有生命的，有人干预的和没有人干预的，地上的和天上的。他还区分了各种不同的运动：事物生死的实质变化（烧木头产生的火），事物生长或萎缩的量的变化，某个特性向另外一个特性的转变（绿叶变成棕色），局部运动或非局部运动。

亚里士多德看待这些变化时采用的是生物学的视角。他将世界看作一个宇宙生态系统，这个系统中包含了各种层次的组织。在该系统中，运动几乎从来都不是随机混沌的，而是从一种状态变为另一种状态的过程。在该过程中，物体是以潜在形式（形式原理）而不是实际形式存在的。人由器官组成，而许多人构成一种状态，各种层次的组织相互交错，因此任何事件都是由各种不同类型的原因导致的复杂网络形成的。

亚里士多德是在一组有着重要差异的框架下来理解宇宙生态系统的。比如，他区分了两种不同的运动：自发运动和受迫运动。自发运动就是物体在原处发生变化：橡子长成橡树或者鸡蛋孵出小鸡。此时变化把物质中的内在原理变成了现实。而受迫运动是由外部施加的外力导致的变化，比如人砍倒橡树盖房，或者拿鸡蛋当食物。

亚里士多德认为运动与运动发生的地点也是有关系的。在地面上，物质是由土、气、火、水，以及不能连续运动，只能间歇运动的物体组成的。而在天上，物体是由一种不变的称为“以太”的物质组成的。以太不停地作圆周运动。人们如今会发现上面的叙述是错误的，由此不难看出从亚里士多德的时代到今天人类已经有了多大的进步以及人类的视野发生了怎样的变化。亚里士多德的思想都是基于理性思辨、逻辑推断和仔细观察的。几百年来，除圆周运动之外，希腊和世界上其他地方的天文学家没有目睹到天体行为的任何变化，也没有发现任何新事物[2]。

于是亚里士多德认为只有圆周运动是不会停止的，并且只有一种地上没有的特殊物质（命名为以太）不会发生变化[3]。天体的运动是由“不运动的原动力”启动的，它使天体不断运动。这个“原动力”就是亚里士多德所谓的上帝相似物，尽管它不属于人类，跟人也没有任何“关系”（今天21世纪人们的叫法）。通过各种媒介物，天体把运动传递给了地球。因此宇宙生态系统中的所有运动，不管多么微小，都是通过宇宙第一原理间接联系在一起的。只有在这个前提下，人们最终才能理解运动。

在亚里士多德讨论我们所谓的运动的2500年之后的今天，作为局外人的我们在解读他的思想时，要注意避免加入自己的假设。亚里士多德所说的局部运动，背景常常是马拉车或者造船工人推船。这得先有了目标、计划和设计组成的复杂网络之后，才能实现。而这当中，局部运动只是一个方面。亚里士多德看到这点之后，并没有脱离整个背景，去为局部运动提出假设并为之辩护。相反，他只是笼统地说，如果要解释其他现象，需要怎样去做。再进一步，在这些类型的事件中，加速度几乎没有任何作用，而经验法则却频频出场。例如“使物体抵抗阻力在一定时间移动了一定距离的力，可以使相同的物体在一半的时间之内移动一半的距离”。亚里士多德还提出：“如果一个物体在一定时间内下落了一定的距离，那么重量等于该物体两倍的物体只需一半的时间就能下落相等的距离。”[4]这个蹩脚的断言在2000年之后被自由落体实验所击溃。

我们很难站在亚里士多德的角度上去看世界。因为有了匀速和加速度等熟悉的概念，人们对运动的理解完全是基于一种定量的方式。如今，人们处在一个技术手段高度丰富的环境中（有电子钟表和速度计等仪器），对于仪器的日常体验依赖于概念和仪器本身。人们对运动的理解不是本能，而是后天获得的“第二本能”。亚里士多德和与他同时代的人的体验与今天的人们不同。他们既没有实验仪器，也没有能对运动进行测量和分析的数学基础（人们对此并没有迫切的要求，也就不会急于去建立它）。他们发现用形式和目的（而不是运动发生的快慢）来理解运动也

是可行的。

亚里士多德和与他同时代的人对公式 $F=ma$ 中的关键部分并不熟悉。他的速率或“快慢”概念仅仅是某个物体在一定的时间内比其他物体经过了更长或更短的距离——我们称为“平均速率”或者“整体速率”，而不是瞬时速率（特定时刻的速率）[5]。他的加速度概念就是物体在接近天然位置时，运动速度会更快[6]。他没有质量的概念：质量是对推力的抵抗程度，不等于重量。他也没有力（dynamis）的定量概念以及力的单位。力是运动的量度。

但是，把自然界看作一个巨大的、包括各种物质的生态系统却是有意义的。这些物质通过对其他物质施加各种内部排斥力，从而相互影响。万物的运动都有目的，这对于维持宇宙生态系统的不同区域是至关重要的。要理解大自然就要在完美状态下观察自然现象。“完美”的意义是完全呈现之意，比如大树、成人和井然有序的社会。这些现象都有各自的作用。在这种情况下，现象产生的原因和过程是最清楚的。

亚里士多德喜欢说的一句话是：智者总是尽可能准确地去寻求研究对象的信息。他用恰如其分的准确字眼对观察到的现象进行描述。与理解自然界中的运动有关系的，是形式、物质和目的三者在将可能性变为现实的过程中所起的作用。人们最终还得回到“不动的原动力”，是它用爱把外部星球、月球和地球联系了起来。

超越亚里士多德

亚里士多德的宇宙图景对西方文明有着巨大的影响。他的思想经他本人创立的学校——“学园”的学生和著作的注释者流传下来。一开始注释者只是希腊人。在 9 世纪到 12 世纪，阿拉伯人也参与了进来。西方学者又通过阿拉伯人知道了亚里士多德。

但是亚里士多德的宇宙图景在某些方面并不令人满意，甚至连他本人也不甚满意。比方说，他似乎对于抛射物和制陶器用的转盘在起始推

力作用下，就能一直转下去的现象感到困惑。如果原动力必须要始终与物体保持接触，物体才能运动下去的话，那么石头和箭为什么不会在被抛出或射出后立刻落到地面上呢？他考虑了两种可能性：一种可能性是原动力（抛掷者或弓）在抛射物（石头或箭）的附近注入了一种介质（空气），从而使物体能始终保持运动[7]；另一种解释叫作逆环境学说，即抛射物前部的空气会跑到抛射物的后部，推动它前进[8]。不过亚里士多德本人对上述两种解释仍不满意。

之后的思想家对亚里士多德的上述解释和他对运动的解释也不满意。反对的声音有些是合乎逻辑的，有些是从经验出发，还有些是两者兼而有之。于是人们开始对亚里士多德的概念进行讨论、研究和修改，引入新的概念。经过几千年的时间，人们才慢慢将注意力集中到了运动的各个层面，最终导致了方程 $F=ma$ 的出现。在这个过程中，人们并没有看到 $F=ma$ 中的任何部分。但这其中的每一步都是必不可少的。下面介绍一下过程中的一些发展阶段。

公元前 3 世纪，小亚细亚拉姆普撒科斯的希腊人斯特拉图（公元前 340—前 268 年）于公元前 287 年掌管了学园，发展和拓展了亚里士多德的思想，写出一本颇具影响力的书——《论运动》（*On Motion*）。斯特拉图发现，要使思想符合逻辑和经验，必须要修改，甚至抛弃亚里士多德的某些观点。其中一个要修改的就是认为自然运动有两种（上和下）的看法。斯特拉图认为所有的物体都是要落到地球的中心的；火和烟等比较轻的物体之所以能上升，是因为受到了更重的物体的替换和排挤。他还受到两个现象的困扰，这两个现象表明物体在下落的过程中速度会加快。一个就是雨水从屋檐上流下的时候，一开始是连续的，但后来就破碎成为水滴。如果不是因为水滴的速度加快了，这种现象是不会发生的[9]。另一个现象是，从空中扔下一块石头，石头起始时的位置越高，对地面的冲击力越大。为什么会这样？石头的重量并没有变大。斯特拉图的结论是石头的速度一定变大了，也就是说落体“用最短的时间经过了

最后一段轨迹”。这是加速度概念的雏形，比亚里士多德的运动观点要复杂得多。

公元 6 世纪，约翰·菲洛波努斯（John Philoponus，人称“工作狂”，约 490—570 年）进一步对亚里士多德的运动观进行了修改。他从逻辑的观点出发，指出真空中可以有运动（亚里士多德之前认为是不可能的）。他认为速度是由力与阻力的相对大小决定的，从而解释了力与阻力大小相等时会发生什么现象。菲洛波努斯是已知的用不同重量的物体进行自由落体实验的第一人。他发现这些物体的下落速度基本是相同的。直到 1000 年之后，伽利略才得出了相同的结论。但是菲洛波努斯在对亚里士多德思想的修改中创新最大、意义最深远的是抛射物的运动。他反对逆环境学说：如果原动力能与抛射物后面的空气就运动进行通信，那么人们为什么不用手搅动石头和箭后部的气体，让它们保持飞行状态呢？菲洛波努斯提出，扔石头的时候，手所施加的力并不是作用在空气上，而是作用在石头本身上。“施加的外力”从抛射物的内部使抛射物继续运动，但该力在克服介质阻力和自然的下落力的过程中逐渐被消耗。当这个力被完全耗尽后，自然运动最终取而代之，或者说石头就落到地面上来。这种观点仍以亚里士多德的思想为基础，因为它假定物体不会发生运动，除非通过某种形式的接触，例如落体的重量和手所施加的外力等。不同之处在于，相对运动的物体而言，接触的方式可以是内部的。基于此，菲洛波努斯及其追随者对世界有了不同的理解。他们无需再去区分自然运动和受迫运动，也无需把天地分开。上帝创造了天上的天体之后，便用外力保证它们的运行。天上不存在阻碍运动的介质。菲洛波努斯的思想大大启发了学者们。他们在理解运动时，开始把注意力集中到起点（运动的原因），而不是终点（运动的终点或目标，无论在天上还是地下）上。

菲洛波努斯对亚里士多德的修正，尤其是外力的修正，影响了伊斯兰学者。这些人中包括西班牙伊斯兰教神学家 Ibn Bajja（一般称其为阿文帕塞，约 1095—1138 年）、西班牙伊斯兰教神学家 Ibn Rushd（阿威罗

伊，此人反对菲洛波努斯的观点，1126—1198 年）和波斯伊斯兰教神学家 Ibn Sina（阿维森纳，980—1037 年）。Ibn Sina 将菲洛波努斯外力的思想翻译成阿拉伯语，书名为《疯狂的倾向》（*mail qasrī* 即 *Violent Inclination*），较重的物体下落时相对较轻的物体速度增加得更多。这也是为什么抛出的石块飞行的距离比一把草或者一根羽毛要远。阿拉伯人还考虑了一些亚里士多德的解释不能令人满意的情形：如果在地球中间钻一条隧道，往里面丢一个石块会发生什么现象？粘在箭前部的线在箭飞行时是向前的吗？Ibn Sina 的思想比菲洛波努斯的更加清晰。运动的关键并不在于形式和目的因，而在于效率和物质因。

这种注意力的转移在让・布里丹（John Buridan，约 1300—1358 年）的作品中得到了很明显的体现。布里丹进一步发展了菲洛波努斯和 Ibn Sina 的思想，给出了外力的学术名称：*原动力*。这一说法直到近代才广为人知。与外力不同的是，原动力是永恒的，不会耗尽；物体如果要失去原动力，就只能把它转移到其他物体上[10]。原动力听起来有点像我们所说的惯性概念，但它又不是概念，只是“因”。因此，新的框架仍然是亚里士多德式的，它仍保留了自然运动和受迫运动之间的差异，认为石头和箭等抛射物的运动是由于不断受到“因”的作用。但是“因”（原动力）是在物体内部起作用的。这样一来，困扰亚里士多德的几个关键问题就迎刃而解了，例如抛射物的运动问题和物体如何下落的问题。石头能够飞行是因为抛石者将原动力传递给了石头，而不是传递给了介质；下落的物体在落下时也获得了原动力。这就解释了为什么落体的速度会增加。原动力的思想使人们形成了质量概念（物体抵抗运动的能力，与重力不同）的雏形。炮弹具有的原动力比木头要大。这也解释了为什么天体不需要神的介入就能一直运行下去：因为天上没有阻力。上帝创造出天体之后，就赋予他们原动力。所以上帝在创世的第七天就可以坐下来休息，而创造物并不会停止运动。这样看来，上帝的作用就与亚里士多德的看法正好是相反的：他并不是保持天体运动的终极因，只是使天

体发生运动的动力因[11]。

接下来的 300 年间，学者们开始采用原动力的思想来理解和解释运动。此时，区分自然运动和受迫运动性质不同的必要性已经减少，不过仍未消除。不同物质之间的不同以及天地的不同，情形是类似的。于是，人们针对力提出了新的概念，比如撞击力（立即产生效果的力，例如球拍对球的撞击）以及通过一定距离持续施加的力（地球对物体的引力）。这些都有助于质量概念的形成。质量是抵抗力的物体的内部物质密度，它与重量有关，但又不是重量。另一方面，学者们也开始在哲学家泰勒所谓的“固有框架”下，对运动本身进行研究，并在不考虑宇宙其他部分的目的、计划和设计的前提下，对运动的某些（并非所有）方面进行研究。这时，他们开始看到所谓的物理学与形而上学的分离。科学世界与日常生活世界开始脱离。

与此同时，人们开始用新的方式将数学应用于宇宙研究当中。数字用于事务处理已经有几百年的历史，但学者们仍在不断提出新的工具，来深化和拓展数字的用途。牛津大学莫顿学院的托马斯·布雷德沃丁（Thomas Bradwardine，约 1300—1349 年）就是其中一位。乔叟在《坎特伯里故事集》（*Canterbury Tales*）中曾提到此人。此人在当时颇为有名，后来成为坎特伯里的大主教。布雷德沃丁为了使数学能够处理速度、瞬时速度（某一时刻的速度，而平均速度是一定时间内速度的平均值）、匀速、匀加速和变加速等概念，提出了数学的基础架构[12]。他用数学的形式，彻底改写了亚里士多德、菲洛波努斯和阿威罗伊等人的观点，指出了他们观点的局限性，并提出了新的运动定律。布雷德沃丁的工作随后又经过尼古拉斯·奥斯姆（Nicholas Oresme）的改进。奥斯姆提出可用数字表示任何连续变化的量，例如运动和热等。他还说：你就“想象”自己在测量几何平面就行了[13]。

菲洛波努斯及其追随者（被称为牛津“计算器”）、奥斯姆和同时代的其他人并不是实验学家，但他们却为后来的实验学家提供了一个复杂

的数学架构，为数字在自然界中的广泛应用铺平了道路。后人在使用数字的时候变得很自然，不会再去“想象”几何平面上的测量。16 世纪晚期和 17 世纪早期，数字被广泛用于各种现象的定量中。威廉·哈维（William Harvey，1578—1657 年）定量得出了心脏的泵血机制；桑塔雷欧·桑塔雷欧（Santorio Santorio，1561—1636 年）定量得出了人体的食物摄入和排出量[14]。这些都极大地影响了人们对运动的理解。许多早期的思想家，如圣·托马斯·阿奎那（St. Thomas Aquinas），已经通过以其他形式（苹果的红、人的善心）表现出来的物体（苹果、人）的参与程度的增减来理解各种变化了。但是随着数学的发展，人们开始把发生的所有变化都看作加减法。这与增减线段以改变原有线段长度的做法类似。

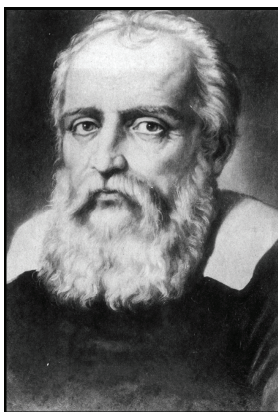
另外，还有一些事件也使人们产生了对亚里士多德框架的不满意，为公式 $F=ma$ 的提出进一步铺平了道路。1572 年和 1604 年相继出现了两次超新星爆发。天文学家指出，爆发并非发生在地球附近，而是发生在天上。很显然，天上的物体变化规律与地上的物体是一样的。1609 年，伽利略（Galileo Galilei，1564—1642 年）用望远镜观察天空，提出天地之间的类似之处比之前想象的要多。这一系列事件推动了宇宙物理学的出现。其他方面的一些发展也使人们对力的理解发生了改变。1600 年，伊丽莎白女王的御医威廉·吉尔伯特（William Gilbert）写了一本磁学专著。这是近代科学首次对磁学进行论述。吉尔伯特认为，磁铁有磁性是因为它向外发射磁力线。他认为地球本身就是一块巨大的磁铁，不断向空间中发射磁力线，且磁场强度随着距离的变化而变化。这有助于人们理解不需接触就能产生力的思想。但这种思想与亚里士多德的思想却迥然不同。约翰尼斯·开普勒（Johannes Kepler，1571—1630 年）在他出版的两本书《新天文学》（*New Astronomy*, 1609 年）和《世界的和谐》（*Harmonies of the World*, 1619 年）中，给出了行星轨道的三定律，堪称宇宙的数学脚本。开普勒认为行星的运动是上帝的意旨。上帝之所以让

行星运动时遵循这些定律，是因为他从定律中发现了美。宇宙中的数学乃是宇宙的脚本和终极因。

伽利略的思想更加激进，认为我们不仅能够读懂宇宙的数学脚本，而且别无他法，只能如此。所有其他的“因”都是要忘掉的。他这样写道：“自然之书是用数字写成的。”那种要去寻找终极因的幻想是没有意义的。为增加书的可读性，伽利略想出了一个聪明的假想实验：考虑在一个没有阻力的无限大的平面上会发生什么现象。为了理解平面上物体的运动，他接着用分段实验研究了钟摆和斜面上的滚球等。这些实验引入了与亚里士多德完全不同的时空观。亚里士多德将空间看成是边界，而伽利略则将空间看作是具有某些几何属性的容器。要理解运动，就得知道物体在一定时间内（以心跳或水滴数为单位）经过了多远的距离（伽利略采用“腕尺”作为计量单位）。在该过程中，伽利略发现了著名的落体定律。它采用比例形式对定律加以表述。表述的形式与今天人们熟悉的落体定律形式不同。就像布雷德沃丁改写前人的思想一样，人们也对落体定律进行了重写。落体定律现在的形式是 $d=at^2/2$。这是第一个真正意义上的自然数学定律，也是第一次用与后来的公式 $F=ma$ 相同的科学语言写成的。伽利略在分析时，将炮弹、大理石和钟摆等的运动分成两部分：“匀速运动”（向一侧）和“加速运动”（向下）。

然而，伽利略没有提出公式 $F=ma$ 中涉及的内容。他依旧受亚里士多德的传统思想的影响，对自然运动和受迫运动加以区分，如自由落体、“受迫”推力或外界施加的推力等。伽利略的思维方式倾向于后者。例如，他并不认为落体因为受到了力的作用，才导致了速度的增加。这阻碍了力和力在运动中的作用的总体概念在伽利略头脑中的形成。与此伴随而来的是用词的不确定——他不知道什么是力，经常采用原动力、力矩、能量和力（源自拉丁语 fortis，意思是强的、强大的）等近义词来表示力[15]。他提到力的时候，一般指的不是所谓的连续力，而是瞬时力（一个物体击打另一个物体，如撞球杆撞击球，锤子撞击钉子），或叠加

在一起的一系列瞬时力。伽利略有着模糊的质量概念，认为它是物体抵抗力的一种属性，是物质的密度。它与重力有关，又不等于重力。即使没有重力，质量也还是存在着的。在现代人看来，伽利略的很多思想听上去都是很奇怪的，例如他曾说圆周运动有助于整体中部分的有序排列，以及直线运动表明物体不处于自然状态，且正在重新回到自然状态。科学历史学家理查德·韦斯特福尔（Richard Westfall）这样评价伽利略的自然观："这些元素来自完全不同的世界观，而他不仅把这些不同的元素糅合到了一起，而且还能泰然处之。"[16]



伽利略（1564—1642 年）

牛顿

但是，所有的一切在牛顿（1642—1727 年）的《自然哲学的数学原理》（*Principia*，以下简称《原理》）中却是那么清楚、系统。牛顿本人从伽利略等前人的思想中受益良多，针对连续力和瞬间力提出了更一般的定量概念，定量描述了物体运动的变化。在《原理》一书中，运动的变化不是从物体内部来解释的，而是完全是用外部施加的力解释的。这是一种理解运动的新方式。它不考虑原因，只考虑过程。



艾萨克 · 牛顿（1642—1727 年）

当然，这也只是在伽利略和与牛顿同时代的莱布尼茨和笛卡儿的基础上迈出的有限一步。可正是这重要的一步造就了它的不朽。它改写了

自然本体论：本体论是人们对所看到的世界加以解释的最基本单元。与牛顿同时代的大部分人认为这些基本单元就是物体本身。物体通过某种机制，产生各种变化，对其他物体产生影响。牛顿的不同之处在于，他认为只有通过改变物体运动的力，才能解释运动。运动本体论的 3 个基本元素是：力、质量（抵抗力）和加速度（运动的变化）。这三个量都是定量的、可测的。

牛顿出版的《原理》是科学界最重要的一本专著，其重要性不亚于欧几里得的《几何原本》。《原理》从不证自明的公理出发，推导出了书中所有内容。《原理》共分三册（前两本是“物体运动”，第三本是“世界体系”），最前边是前言、8 个定义和 1 组“公理，或称运动定律”。我们可以看到，牛顿在前人的基础上，在定义中（有时是很笨拙地）给出了公式 $F=ma$ 中的元素，特别是力的思想。定义 1 是有关质量或者物质的量的；定义 2 是有关运动量的。后面的定义是关于各种形式的力，里面存在表达不清甚至令人费解的情况。例如，定义 3 大概想说“内力”，可实际上说的却是用推力描述所谓的惯性。一位研究者称，最好把这些定义理解为“对伽利略之前的力学的妥协”[17]。关键是定义 4，它用现代术语定义了“外力”，即“施加到物体上的改变其静止或匀速直线运动状态的力”。于是牛顿把前人提出的力的概念一般化了，从瞬时力扩展到了连续力。科学史学家伯纳德·科恩（I. Bernard Cohen）认为，这一切全凭牛顿本人的直觉，这一步“靠的既不是严密的逻辑推理，也不是实验”[18]。正如牛顿后来写的那样：“这个概念完全是数学上的，我并没有考虑物理上的原因和力的位置”[19]。剩下的 4 个定义涉及力的其他方面。牛顿在注释中警告说，“虽然所有人都知道时间、空间、地点和运动”，并且通过感官体验赋予了它们通俗的意义，但是他还是要对这些量进行专门的定义，描述他所谓的永远流逝、不会停歇的“绝对时间”，以及均一静止的“绝对空间”。之后牛顿又区分了绝对运动和相对运动。“绝对运动是物体从某个绝对位置到另外一个绝对位置的位置变化；而相对运

动是从某个相对位置到另一个相对位置的位置变化。”[20]这样一来，船上水手的运动就是 3 种运动的叠加：水手相对于船的运动、船相对于地球的运动和地球相对于绝对空间的运动。

有了这些定义，接下来就是“公理（运动定律）”了。第一定律是关于惯性的，“物体倾向于保持静止或者匀速直线运动状态，除非有外力改变这种状态”。人们对此是这样描述的：“这是 17 世纪把神灵从世界中驱除出去的最伟大的因素，人们从此可以像理解时钟的部件那样来理解世界。”[21]牛顿第二定律的内容是：“运动的变化量与施加的力成正比，且变化的方向沿着力的直线方向。”不过他并没有用方程表达出来。我们现在的写法是 $F \propto ma$ 或 $F=ma$，因为速度的变化就是加速度。伦纳德・欧拉（Leonhard Euler）是用方程表示牛顿第二定律的第一人。该方程是他在一个世纪之后提出来的。牛顿第三定律的内容是：“有作用力就有反作用力。两者大小相等，方向相反。”

牛顿为什么要把上面三条称为公理或定律呢？“公理”一词用于表示所讲的是先天（通过定义或先验）就成立的；而“定律”指的是通过日常生活的经验发现的。实际上，牛顿的公理−定理中二者兼而有之，它以运动形式的定义，和一定条件下从实验室的测量结果获得的经验事实为基础。这些公理−定理描绘出了客观知识的轮廓和不同条件下的不变量。棒球等物体有重量，但到了月球或者火星上重量就会变化。为什么会这样？棒球还是那个棒球。这里，客观不变的是质量，而不是重量。质量是与引力无关的。与此类似，如果将棒球抛出，它就会以一定的速度飞行，但是用同样大小的力却没法使炮弹也以同样的速度飞行。物理学的规律是不变的：在任何情况下，力、质量和加速度三者之间的关系都是绝对不变的。按现在科学家的说法，牛顿第二定律说的就是“不变性”。因此科恩认为“第二运动定律的所有实际应用都是以牛顿的物理思想体系为核心的。”[22]

在《原理》一书中，牛顿对伽利略的无限平面上的无阻力假想实验

加以推广，提出了完整抽象的世界全景图。所涉及的量只有力、质量和速度，不涉及人的目的和终极因。因此，在这种条件下看待运动的某些方面要比实际情况更加清晰。亚里士多德认为不可能也没必要做到这一点，他认为运动交织在一起，从外部是无法理解的。（牛顿在第三定律的注释中，通过分析马拉石头来解释涉及的力。读者可以想见亚里士多德对此一定是非常困惑的。他可能会问：是谁用绳子把马系到石头上的？为什么？这么做的目的是什么？）而伽利略对此则是远远地观望。牛顿的体系虽然简陋，却蕴含着美和力量。我们不能延续亚里士多德把世界看成宇宙生态系统的看法，认为不同的物体在不同的区域产生的作用不同。世界更像一张台球桌，其中的所有空间和方向都是类似的。所有的事件都是运动，所有运动的变化都是因为同一种基本类型的物体被施加了同一种基本类型的力。运动涉及的是空间上的变化，不涉及到达、实现或加强。所有的吊灯、吊环和秋千都是摆，所有的运动和舞蹈都能用公式 $F=ma$ 表示，所有的球都是弹性的，所有的平面都是无限延伸的。你尽可以在空间和时间中运动，不管是在火车、飞机、过山车还是自行车上。所谓的“平移不变性”说的就是这点，即定律始终保持不变。按照牛顿的理论，要理解运动，首先要知道位置、速度和质量，然后是力[23]。

由此可见，$F=ma$ 既可以是定义，也可以是可观察的经验事实。说它是定义，是因为它就像抽象世界的“经纬线”；说它是事实，是说如果用正确的概念、假设和测量技术把它与现实世界联系起来，就能定量描述实验室条件下观察到的数值之间的关系。理论是往返于抽象世界的理想值与现实世界的实际值之间的工具。在科学家以外的人看来，抽象世界可能有些奇怪，好像是往自然界中任意加入了某些东西，是虚拟世界。这种虚拟固然可能很有效，但还是属于发明的范畴。而在科学家看来，情况正好相反。他们都受过良好的训练，惯于用概念、程序和测量手段把抽象世界和现实世界联系起来。他们会忘记抽象世界的抽象之处，频

繁往返于抽象和现实之间，好像抽象世界本身也是现实世界的一部分。

最后，$F=ma$ 的出现表明，人类对运动的理解从最初到现在已经走过了很长的历程。我们仰望天空，俯视大地，不再区分自然运动和受迫运动之间的不同，不再区分自然位置和非自然位置之间的差异，也不再区分力的类型之间的不同。所有的物体都遵循相同的规律。如果有无法用规律解释的现象，那是因为人们还不知道用什么规律去描述该现象。运动不是一个动作，而是一种状态。运动只有一种。圆周运动可以用运动和力组合起来加以解释。亚里士多德认为运动在天上和地下是不一样的。果真如此的话，是因为天上的阻力和地下的阻力不同。自然是由力、加速度和运动之间的时空关系决定的。人们可以从理论上推出这种关系。我们可以把宇宙理解成由定量规律构成的网络体系。这些定律可由实验发现。

伏尔泰在 1732 年写道："牛顿就是我们的哥伦布，他把我们带进一个崭新的世界中。"[24]但是这也是一个奇特的世界，它既不同于新大陆的发现，也不是经由仪器揭示出来的。它与显微镜下的微小世界或者望远镜里的遥远天体还不一样。牛顿的奇特世界源自现实世界，却又不是现实世界。人在牛顿的世界中是无法生活下去的。包括科学家在内的所有人都是生活在哲学家们所谓的"日常生活的世界"中，与目的、欲望和目标打交道。人们生活在一个亚里士多德式的世界中。牛顿发现的是抽象世界，它会改变人们对事物的看法和看待事物的角度。从外部看，牛顿的世界就像是鱼缸，它与现实世界相隔离。但是透过牛顿抽象世界的事件，人们就可以对现实世界有更多的了解。正如维尔泽克所写的："方程 $F=ma$ 是牛顿抽象世界的'灵魂'，它定义了世界的结构和世界中发生的事件的主要部分。"

方程 $F=ma$ 之所以看上去不是那么直接，就是因为上述原因。人们在学习它的时候所真正学到的，比最初想象的要多。如今，我们又把前人一步步得到方程 $F=ma$ 的"旅程"继承下来。



茶　歇

自　然　之　书

天地万物这部大书是用哲学语言写成的。它始终向人类敞开。要读懂这部书，就得先理解书中采用的语言，读懂书中的字母。写成这部书的语言是数学，语言符号有三角形、圆形和其他几何形状。没有它们，就无法理解书中的字，人类就只能在黑暗的迷宫中徘徊。

——伽利略，《试金者》（*The Assayer*）

1623 年，伽利略提出了一个著名的见解，至今仍为科学家们广泛引用。他写道：大自然是一本用“数学语言”写成的书。人们如果不懂这种语言，就注定要在自然的“黑暗迷宫”中徘徊。

与所有的比喻一样，伽利略的也是亦真亦假。这个比喻确有见地，但如果只从字面上去理解，不免会产生误导。它抓住了人们的一个观念：认为自然事实是强加在人身上的，它们是描述真理的数学语言的基础，早已深深印刻在世界之中。但伽利略从特定目的出发，又对见解进行了修正。他把自己的见解从历史背景中取出，放到现在的背景下，具有危险的迷惑性。

第一个提出自然之书想法的人并不是伽利略。几个世纪以来，人们一直都把世界就是两本书的想法作为宗教教义的一部分接受下来。第一本书就是大自然，它全部由揭示深层意义的符号组成。这种深层意义是按圣经解读出来的。第二本书中是终极意义或者自然特征的句法。要理解世界，就要去读这两本书，在观察到的世界和所读的圣经之间不断往复。正如皮特·哈里森（Peter Harrison）在《圣经、新教和自然科学的

崛起》（*The Bible, Protestantism, and the Rise of Natural Science*）一书中所指出的，人们曾经认为读圣经就是研究自然，而人们现在常把读圣经这样严肃的事情与文本主义（literalism）和反科学行为等同，这种做法是不对的。

但到了文艺复兴时期，学者们开始敏锐地感觉到，自然事实并不都是能够轻易看出来的。相反，真理常常是隐藏在自然深处，要找到它们需要经过特殊训练。同时，新教改革（Protestant Reformation）使人们对文字的理解发生了变化。人们开始更加重视文字中蕴含事实的准确性和完备性，而不是文字的象征性和寓意。

伽利略在科学革命和宗教变革的基础上，按照自己的需要将"两本书"进行了合并，改变了它们原有的意义。

1623 年是伽利略身处困境的一年。其实早在 10 年前，他的学生在比萨宫廷上与别人讨论他的工作时，麻烦就已经找上门了。当时有人发现伽利略的科学观点与圣经有明显的冲突，特别是在关于地球运动的问题上。另一方面，因为同样的原因，当局威胁要将伽利略的科学伙伴哥白尼的《天体运行论》一书列为禁书。出于对自己和其他科学家的担忧，伽利略写了一封信给大公夫人克里斯蒂娜，向她说明了科学与圣经之间的关系。伽利略在信中借用上帝的传统形象，说明自己是通过上帝的两本书了解了人性。这两本书就是自然和圣经。他认为这两本书都表达出了永恒的真理，而且写出这两本书的都是上帝，因而它们是可以和谐共处的。上帝只不过是用两种不同的方式说了同一件事情。

然而伽利略所打造出的世界新形象却有着爆炸性的作用。他坚持认为自然之书不是用一般的文字写成的，这些文字与圣经中的文字、亚里士多德的文字和其他任何作者的文字都有着根本的不同。伽利略说："为了让每个人都能理解，并且在只考虑意义的情况下，提及看似与绝对真理不同的许多事物，圣经需要采用这样一种语言。但另一方面，自然又是不可改变的、永恒的；它不会超越自身的规律，也丝毫不在意人们能否理解它深奥的运行规律。"[25]

伽利略的意见似乎说服了克里斯蒂娜，但却没能说动权威。1616 年《天体运行论》被列为禁书。随后的 1619 年，开普勒根据哥白尼的天文学论著《哥白尼天文学概要》（*Epitome*）写出的教科书和伽利略本人也遭到了攻击。作为对攻击回应的一部分，伽利略写出了《试金者》（*The Assayer*），其中有这样一段著名的话："要读懂天地万物这部大书，就得先理解书中采用的语言，读懂书中的字母……语言是数学……"换句话说，在数学和物理方面有造诣的人能够体会得到伽利略那种一般人体会不到的成就。

伽利略为世界选择形象时是很仔细的。这个形象深深扎根于西方的形而上学和神学。第一，它的思想是传统的，认为上帝会在世界中显示他的力量、荣耀和真理。第二，它还依赖于另一个传统思想，即圣经不能与显而易见的逻辑和人的感官相矛盾。最后，它借用了一个古老的比喻——把自然比喻为一本书。因此，伽利略的思想有着坚实的神学基础。

但实际上，连伽利略自己都没有意识到，他颠覆了传统的世界观。自然之书中的世界形象与之前世界的形象在某些方面是对立的：自然特征有着自身的完备意义。人们不需要圣经作为象征来理解自然；对自然的研究已经成为一种独立的活动，最好是由独立、专业的学者阶层进行。自然之书现在已经变成了用技术语言写成的蓝图原文，而圣经则是用通俗语言写成的手册。

伽利略于是就用这种世界形象为自己和所有科学家辩护，认为科学家的权威性与神职人员一样具有权威性，不容置疑。哈里森这样写道："自然之书和解释它的哲学家们接替了以前的圣事和教师的角色。"

但是，自然之书的形象现在依然萦绕在人们心中。这其中有一个原因就是它暗示了某个终极的一致性真理的存在——一个"终极理论"。虽然许多科学家对此深信不疑，可它终究只是一种想象。随着概念和技术的发展，人们极有可能在自然界中发现更多的理论。而且世界的形象也暗示出自然之书的"文字"源于神授。有种观点认为世界就是一位具有

超能力的作者写成的全部著作。从这种观点又发展出了认为世界是一位聪明的设计师设计的杰作的观点。当代的科学社会学家们屈从于上述意义，把科学家看作是“从事”或者“想要去从事”与教士类似的工作。

我们从伽利略身上学到的最重要一点是，在科学讨论中，要不断提出比喻，并对其加以修正。

注 释

[1] 个人的主观经验是以自我行动为中心的。它被投射到非人事物上并成为事物的属性之一。这也表明了哲学家 Maxine Sheets-Johnstone 所谓的“活体即是原型的来源”。她指出该过程是早期科学概念形成的关键。类比思考，亦即用熟悉的事物去理解不熟悉的事物就是一个经典的例子。这个例子中最值得注意的就是，熟悉的事物在生命的触觉–感觉体验中存在着基础，因此，它也是有物质来源的。关于早期力学思想和宗教思想之间的联系，参见马克思・詹摩尔的 *Concepts of Force*: *A Study of the Foundations of Dynamics*，剑桥：哈佛大学出版社，1957 年，第 2 章。

[2] 几乎所有亚里士多德的现代版本中都会标出在 1831 出版的标准版中的页数。引用的时候，标准的做法是引用亚里士多德所著书籍名和章数，有时还会加上行数。由于亚里士多德书籍的版本和译本众多，上述做法是很有效的。该条引用是《论天》（*On the Heavens*），第 1 册，第 3 章，第 207b 行 13~17。

[3] 他推断说，天的不变之处在于“永不停息的圆周运动”。这一点不仅可以清楚地推出，也是明显的事实。*Metaphysics*，1072a21。

[4] 全部的规则可以在《物理学》（*Physics*），第 VII 册的第 5 章，《物理学》第 VIII 册的第 10 章和《论天》第 I 册的第 1 章中找到。这些规则包括“如果力变为原来的一半，那么同样的物体在相同的时间内所通过的距离也将变为原来的一半”；“如果物体的重量变为原来的一半，那么物体在相同大小的力的作用下通过的距离将变为原来的两倍”；“如果物体受到的阻力变为原来的两倍，那么它通过的距离将变为原来的一半”；“介质的浓度越大，物体在其中下落的速度越慢”；“物体越重，下落的速度越快”。现在人们喜欢用数学公式表述上述规则。后来的学者，站在几千年后的立场上回头看亚里士多德的思想时，将亚里士多德关于运动的叙述改写为：如果保持力

的大小不变，那么在相同的时间内，物体所通过的距离与它受到的阻力成反比；如果保持力的大小和通过的距离不变，那么所需的时间与物体受到的阻力成正比。人们经常还会进一步作数学上的简化，把距离和时间用速度（V）代替，用 F 代表力，R 代表阻力：

$$V \propto F/R$$（速度与力和阻力的商成正比）

以及

$$V \propto W/R$$（速度与重力和阻力的商成正比）

但是上述表达式是误导性的，并没有反映出观察到的事实。亚里士多德知道上述公式可能存在例外，甚至有公式不适用的情形存在。例如，力和速度之间的联系并不是连续变化的——50 个人推动船只移动的距离是 100 个人推动船只移动距离的一半，但是 1 个人却是没法推动船的，也就是说运动距离为 0。但是依照亚里士多德的定律，情况却不是这样的。他相信物体的速度在接近自然位置时会增加，而这一点却未能在他的规则中得以体现。

[5] 亚里士多德没有匀速运动的概念。他对于运动状态，如匀速运动、加速运动和匀加速运动等，并不感兴趣，而更关心运动的物体从何而来，到哪里去。因此，他不会区分瞬时速率和平均速率。对他而言，速率只有区间速率一种，也就是物体通过一段距离所需的时间。他发现有些运动所需的时间比其他的要长。林德伯格写道："把速率看作一个科技术语，并提出速率可以取数值，源自中世纪人们的贡献。"戴维・林德伯格，《西方科学的起源》（*The Beginnings of Western Science: The European Scientific Tradition in Philosophical, Religious, and Institutional Context, 600 BC to AD 1450*，芝加哥：芝加哥大学出版社，1992 年），第 60 页。

[6] 《论天》，第 I 册，第 8 章。

[7] 同上，第 III 册，第 2 章。

[8] 此处亚里士多德的两个观点，见《物理学》的第 4 册和第 8 册。要了解对这两个观点的分析，参见 Marshall Clagett， *The Science of Mechanics in the Middle Ages*，麦迪逊：威斯康星大学出版社，1959 年，第 505~509 页。

[9] 参见 Clagett，*The Science of Mechanics*，第 258~261 页。

[10] "推动力具有永恒的本质。这种本质不同于使抛射物发生运动的局部运动……"引自戴维・林德伯格的《西方科学的起源》，第 303 页。

[11] "没有必要把信息作为天体运动的推动力……我们可以说，上帝创造天体时，就让天体以自己希望的方式运行。因为不存在阻力，原动力就不会损耗或消失，所以一开始赋予天体的原动力就能一直让天体运动下去。"引自

Simon Oliver 的 *Philosophy, God and Motion*，纽约：罗德里奇出版社，2005 年，第 152 页。

[12] 比如，他证明了平均速率定理，即匀加速运动的物体（如汽车在 1 分钟内速度从每小时 0 英里均匀加速到每小时 60 英里）所通过的距离等于物体以平均速度匀速运动时（以速度每小时 30 英里运动 1 分钟）所通过的距离。詹摩尔，*Concepts of Force*，第 66 页。

[13] “对点、线和面，或者它们的属性加以想象是必要的。点和线虽然既看不见，也不存在，但仍有必要把它们虚构出来。”Marshall Clagett 编译，*Nicole Oresme and the Medieval Geometry of Qualities and Motions*，麦迪逊：威斯康星大学出版社，1968 年，第 165 页。

[14] 可参见伯纳德・科恩的 *The Triumph of Numbers: How Counting Shaped Modern Life*，纽约：诺顿出版公司，2005 年。

[15] 关于伽利略对力的使用，参见 Richard Westfall 的 *Force in Newton's Physics: The Science of Dynamics in the Seventeenth Century*，纽约：爱思唯尔，1971 年，第 1 章和附录 A。

[16] Westfall，*Force in Newton's physics*，第 41~42 页。

[17] 詹摩尔，*Concepts of Force*，第 120 页。

[18] 伯纳德・科恩，“Newton's Second Law and the Concept of Force in the *Principia*”，*The Annus Mirabilis of Sir Isaac Newton, 1666-1966*，Robert Palter 编（马萨诸塞，剑桥：麻省理工学院出版社，1971 年），第 171 页。

[19] 牛顿，《自然哲学的数学原理》（*The Principia: Mathematical Principles of Natural Philosophy*），伯纳德・科恩和安妮・惠特曼翻译（伯克利：加利福尼亚大学出版社，1999 年），第 407 页。

[20] 牛顿，《原理》，第 409 页。

[21] Herbert Butterfield，Oliver 在 *Philosophy,God,and Motion* 中引用，第 168 页。

[22] 伯纳德・科恩，“牛顿第二定律”，第 143 页。科恩继续叙述牛顿思想的演变，直到《原理》一书的问世。他说：“第二定律是《原理》一书中牛顿的一个特别引人入胜的成就，因为它向我们揭示了牛顿是如何从碰撞、打击的动力学等现象出发，并对其加以推广，最终形成囊括中心力、万有引力，并推广到连续力的一个颇具争议的新领域的。”（第 160 页）

[23] 正如牛顿在前言中所写的，本书旨在“通过运动现象发现自然中的力，然后再通过这些力去论证其他现象”。例如，他将用力去推断“飞机、彗星、月球和海洋的运动”。《原理》，第 382 页。

[24] Voltaire 给 Pierre-Louis Moreau de Maupertuis 的信，1932 年 10 月，*Voltaire's Correspondence*，第 2 卷，T. Besterman 编（日内瓦：伏尔泰研究院和博物馆，1953 年），第 382 页。

[25] 伽利略，“试金者”，见 *Discoveries and Opinions of Galileo*，纽约：双日出版社，1957 年，第 237~238 页。

3

科学革命的制高点
牛顿万有引力定律

$$F_g = \frac{Gm_1m_2}{r^2}$$

说明：万有引力在物体中普遍存在。两个物体之间万有引力的大小与二者的质量有关，并且和两者中心距离的平方成反比。
发现者：牛顿。
发现时间：1684 ~ 1687 年。

科学革命的亮点是牛顿所发现的万有引力定律。万有引力在所有物体之间都存在，它的大小与两物体质量的乘积成正比，与两物体之间距离的平方成反比。牛顿用一个数学定律就囊括了观测到的宇宙中的物理现象，表明地球上的物理学与天体之间的物理学是相同的、一致的。

——伯纳德·科恩，《科学美国人》（*Scientific American*）

我们都知道，给物体一个推力，物体就会发生移动；我们也知道，从高处将物体抛下（如苹果），它们就会落到地面上。人们对此已经习以为常。但是牛顿却得出了一个发现——万有引力定律。该定律最初发表的时候并不是今天我们所熟悉的方程形式 $F_g = Gm_1m_2/r^2$ ，而是直接用语言描述的。万有引力定律不仅对下落行为进

行了定量，而且指出了其中涉及的关键物理量和这些物理量之间的关系。正如科恩所说，这个关系的出现标志着科学革命的高潮，它将天和地统一起来，并指出二者遵循同样的规律。（这个关系是牛顿在他的《自然哲学的数学原理》一书中提出的，同样在该书中提出的还有他的牛顿第二定律。）然而，它的意义却是极为深远的。正是因为提出了这一关系，牛顿不仅被人们视为科学家、探险家和天才的象征，更被视为人类追求完美和自我实现的象征，前提是亚里士多德的宇宙图景也已逐渐淡出了历史舞台。这是人类全身心投入、集中智慧所能取得的成就。的确，牛顿万有引力定律的发现似乎是与神灵作了一次亲密接触：这可能是人们所能想到的与上帝最为接近的时刻。由此看来，联想其他一些与苹果有关的故事（伊甸园的圣经故事，人类抓住的第一颗知识树上的果实），牛顿由苹果落地而得出该发现并不是什么巧合。

物理学中最大的难题

在亚里士多德的宇宙生态系统中，下落是一种特殊的行为，宇宙中只有特定位置的特定物体才会有这种行为。下落是诸多不同类型的运动和变化中的一种，不仅与潮汐无关，与行星和其他天体的圆周运动也无关。下落是地球上的物体通过内力回到自然位置的自然运动。因此，下落的原因涉及物体的组成、物体在地球上的自然位置以及物体回到自然位置的倾向。在亚里士多德的影响之下，很长一段时间内，人们都认为物体向地球的坠落只是宇宙中几种不同的“引力”和运动之一。同样，人们也相信他的观点：物体下落的速度取决于其本身的重量。毕竟这一点可以由人们的日常生活体验得到证实。在汤姆·斯托帕（Tom Stoppard）的电影《君臣人子小命呜呼》（*Rosencrantz and Guildenstern Are Dead*）中，罗森克兰茨（Rosencrantz）手里拿着一个球和一根羽毛，两者位于同一高度，说：“你可能会觉得这个比这个下落得快（同时松手让两者下落，球先着地）。那你绝对是正确的。”

但有些古代学者却与亚里士多德意见相左，他们提出天地的现象之间存在着各种各样的联系，其中最明显的就是月亮和潮汐。亚里士多德曾努力尝试从力学上对潮汐运动作出解释，他还在解释中引入了风，但其他人认为这种联系从某种意义上来说是直接的。希腊学者波西杜尼斯（Posidonius，约公元前 135—前 51 年）和其他几个古代学者提出了不同于亚里士多德的力的概念。他们认为力渗透在宇宙之中，与任何物质（亚里士多德认为物质是唯一真正存在的事物）都无关，但却将物质联系在一起。这种宇宙力被命名为“交感力”，在希腊语中是“相互感应”[1]的意思。

在古代和中世纪，人们把探讨天体之间、天体和地球之间的物理作用的学问称为“占星学”。我们一定不能把“占星学”与现代社会所见的“迷信”相混淆。“迷信”似乎赋予了信奉者这样一种能力：仅通过人的出生时日就能预先判断出其性格。当然，古代和中世纪的占星学家中也不乏有些冒牌货，专做此类勾当。但是，占星学确有其严肃的一面：它合理地假定宇宙中存在着将某些物体与其他遥远物体联系在一起的物理作用，并坚信这些物理作用是可以研究和描述的。科学史学家戴维·林德伯格（David C. Lindberg）曾说过：“几乎所有的古代哲学家都认为否认这些联系的存在是极其愚蠢的。”[2]占星家们的工作最初对远程力概念的提出发挥了巨大的正面作用。[3]

然而，如何对这些联系（诸如物体的下落）加以解释仍是令人困惑的问题。力究竟是落体本身的一部分，还是独立于落体之外单独存在的呢？抑或是别的什么东西？1504 年，尼古莱多·维利阿斯（Nicoletto Vernias）在写到自由落体时说：“这是物理学上最难的一个问题。”[4]

1543 年，尼古拉斯·哥白尼（Nicolaus Copernius，1473—1543 年）对该问题加以转换，出版了《天体运行论》（*On the Revolutions of the Heavenly Spheres*）一书。哥白尼在书中指出太阳系的中心是太阳，而不是地球。据说这本书在作者临终时才得以出版。虽然该书假设力是由上

帝嵌入在物体之中的，但它却对宇宙力的研究者产生了极大的影响。它说明地球上的物体的重力在宇宙中并不是唯一的，在绕着太阳运行的其他天体上也可能会有重力的存在，月球如此，甚至太阳也如此[5]。所有的物体都有重力。

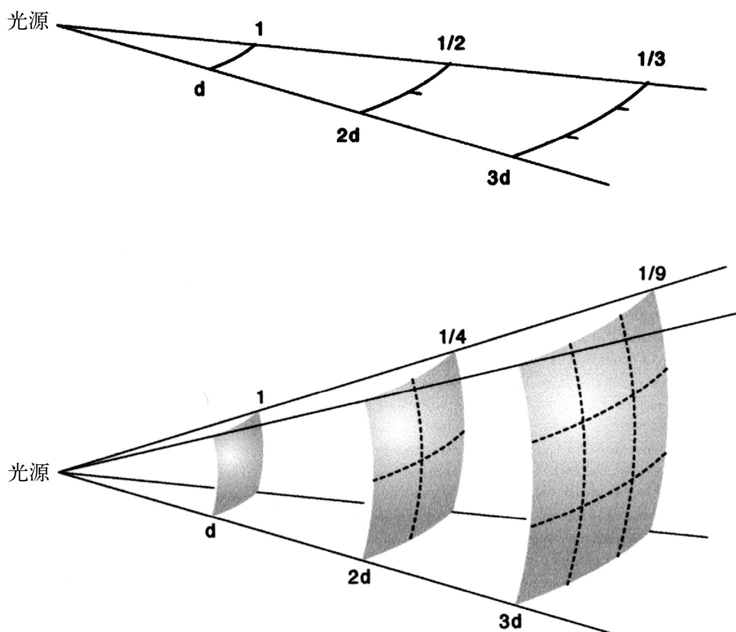
宇宙力思想的另一个里程碑是威廉·吉尔伯特（William Gilbert）的磁力论文《论磁》（*De Magnete*，1600 年）。很明显，磁力的产生是由于地球和各种物质间存在着相互作用。吉尔伯特同时还发现磁力大小随距离大小而变。他还怀疑即便物体处于静止状态，磁力也是存在的。吉尔伯特认为，那种认为静止物体不受磁力影响的想法就像是认为房子由墙壁、屋顶和地板掌管，而不是由住在房中的人掌管那般可笑。

约翰尼斯·开普勒（Johannes Kepler，1571—1630 年）在哥白尼和吉尔伯特的基础上，试着从数学上描述太阳和行星间的宇宙力。因为转而从事数学工作的关系，开普勒的早期神学研究意外中断了。他本人也形成了一种与占星学矛盾的关系。与其他占星学家一样，开普勒热衷于和谐遍布宇宙的思想，认为宇宙依赖于上帝建立起来的和谐。但他对占星学家的方法却抱着轻蔑的态度。这些人在研究中采用的是日常生活式的语言，对专业天文学家所采用的精确数学语言一无所知。开普勒认为：如果没有数学，占星学家就探测不到宇宙中的和谐，因此也就会忽视世界的结构。

开普勒是第一个提出光的强度是依照平方反比定律变化的。平方反比定律叙述的是某个特性随着距离的平方而减小。以光的强度为例，光线是从光源向着各个方向辐射的，即是依据简单的几何学规律辐射的。如果将光源的辐射距离加倍，光线分布的平面面积就会变为原来的 4 倍（光强变小）。如果将距离变为原来的 3 倍，面积就变为原来的 9 倍。

开普勒的天文学思想借鉴了哥白尼和吉尔伯特的思想。他从哥白尼那里借用了太阳系的日心说思想和重力是引力的思想。实际上，开普勒曾基于哥白尼的天文学写过一本著名的七卷本教科书——《哥白尼天文

学导论》(*Epitome to Copernican Astronomy*，1618—1621年)。开普勒从吉尔伯特那里借用了涉及“相互”吸引的力的思想。即便地球是向着石头移动的，石头也还是会向着地球移动。如果两块石头之间相隔一定距离，两者就会相互吸引，且吸引力的大小随着距离的增加而减小。行星与太阳之间的距离越远，引力就越小，行星运动的速度就越慢。但是开普勒给出的结论是太阳与行星之间的引力并不是向着各个方向辐射开去的，而只是延伸到在轨道上运行的行星上。既然引力存在的“目的”只是要“抓住”地球，那么它为什么还要沿着各个方向辐射出去呢？因此，开普勒推出结论：力的大小与行星离太阳的距离成反比，而不是与距离的平方成反比。



在平面上向着一个方向延伸的力将随着离源距离的增大而减小，而沿着各个方向传播开去的力将随距离平方的增大而减小，此即平方反比定律

但是，开普勒在尝试提出控制行星轨道运动的数学关系时却遇到了一个疑惑。按照哥白尼的理论，行星沿着圆形轨道绕太阳公转。这是因为自亚里士多德的时代以来，人们一直认为天体的运行轨道是圆形的。但是，在人们于 1609 年开始将望远镜用于天文学之前，当时的最好数据是由丹麦天文学家第谷 • 布拉赫（Tycho Brahe，1546—1601 年）测得的。开普勒知道有这些数据，对其深信不疑。但这些数据却不能很好地与圆形轨道模型相吻合。两者之间的偏差很小，只差了 8 分①，只比肉眼的分辨率稍大一点，可以说几乎是无关紧要的。开普勒花了 6 年时间尝试将这 8 分的角度考虑进哥白尼的体系中去，但没有成功。

其他人考虑到观测误差和某些未知因素，略去了该偏差。但是开普勒对哥白尼的日心说模型和布拉赫的数据均深信不疑。正因如此，开普勒才提出了一个全新的思想：行星的运行轨道并不是圆形的，而是椭圆形的，太阳位于椭圆的一个焦点上。他还进一步得出结论：虽然行星距离太阳较近时运行速度较快，距离太阳较远时运行速度较慢，但是太阳与行星的连线在相同的时间内扫过的面积是相同的。上述结论就是开普勒三定律中的前两个。开普勒第三定律叙述的是另外一个数学关系：任何两个行星公转周期的平方都与它们离太阳距离的立方成正比。[6]

开普勒发现这些定律不仅很美，而且相互之间还是协调的。他还认为这种美和协调是上帝用这些定律来构建宇宙的根本原因。哲学家伯特（E. A. Burtt）曾说：“这种因果关系的思想本质上是以严格的数学形式对亚里士多德的形式因重新阐述的结果。”[7]开普勒将行星和太阳之间的引力看作不朽的自然力。他曾这样写道：“如果你用‘力’取代‘灵魂’，就可以得出与（我的）天体物理基础一样的理论来。我曾坚信灵魂就是行星运行的动力。但考虑到这个动力随着距离的增加而减小，我推断它是有形的。”[8]开普勒在行星-太阳引力的“灵魂说”与“有形说”之间

① 又称角分、弧分，是量度角度的单位。60 分为 1 度。——编者注

的顺利过渡正好是法国哲学家奥古斯特·孔德（Auguste Comte）所谓的“神学思考与形而上学思考之间的过渡”的一个经典印证。

有形力究竟是什么？人们在整个 17 世纪的大部分时间里对此争论不休。笛卡儿等人认为有形力是纯粹力学的，是微小运动（称为涡旋）在流体样物质（叫作以太）中的产物。整个太阳系就浸没在以太中。[9]伽利略则倾向于完全不去讨论重力的实质，转而集中去测量重力的定量效应。这就是孔德所谓的进入科学思考的一步：给出数字就可以了。

1645 年，法国天文学家伊斯迈尔·布里奥（Ismael Boulliau，1605—1694 年）偶然发现了计算力的大小的公式，可他后来又否定了这个正确的公式。布里奥是科学史上一个颇具魅力的人物。他因思维多变，并设计出准确的天文表而广为人知。他是最先接受开普勒的“行星沿椭圆形轨道运行”思想的天文学家之一，同时也是最后仍严肃看待占星术的天文学家之一。所以，布里奥对开普勒和开普勒对数学的运用进行了抨击。他从占星学出发，强烈反对开普勒的结论——行星的运动是由来自太阳的、与人无关的力控制的，且力的大小随着距离的增加而减小。他对开普勒的荒谬思想嗤之以鼻，说如果存在这样的力，那么这个力就会像光一样，向着各个方向延伸开去。这意味着力的大小将与距离的平方成反比。而这显然是很荒谬的！他不相信上帝会这样来安排。[10]

然而，其他一些科学家认识到行星和太阳之间的力可能的确是向着各个方向延伸开去的。也就是说“平方反比定律”不仅未必荒谬，而且的确可能存在于太阳和行星间的力当中。不过他们认为这种关系可能是离心力和向心力之间“拉锯战”的结果。这个结果就是平方反比定律。他们还猜想从平方反比关系中可以推导出开普勒定律。

罗伯特·胡克（Robert Hooke，1605—1703 年）就是上面这些科学家中的一员。胡克是伦敦皇家学会的实验主管。他于 1674 年提出在地球和其他天体中都存在“指向各自中心的引力”。这个引力不仅作用于天体自身，而且还会作用于“有效球体作用范围”内的所有其他天体。作用

力的大小与天体之间的距离有关。[11]但胡克的数学功底欠佳，无法由此假设计算行星的运动。1679 年，仍在寻找问题答案的胡克写了一封信给身边一位有为的数学家——牛顿。胡克问牛顿对于他的“指向中心天体的吸引力”有什么想法？[12]1680 年 1 月，经过与牛顿的几次通信之后，胡克提出引力的大小是按照平方反比定律变化的。他接着问牛顿，如果实际情况真是这样，能否计算出行星的运行轨道？

无独有偶，1680 年发生的几个事件激发了人们对天体运动的兴趣和研究天体行为的好奇心。一个事件就是天空中出现了一颗很大的彗星。英国著名天文学家艾德蒙 • 哈雷（Edmond Halley，1656—1742 年）对这颗彗星进行了观测。之后在 1682 年和 1684 年分别又出现了另外两颗彗星。现在知道，这两颗彗星就是“哈雷彗星”。此前，人们一直认为彗星是在太阳系中随机出现的外来天体，它们不受定律的约束。这种观点很快就发生了变化。

1684 年 1 月，在伦敦的一家咖啡馆，哈雷、胡克还有科学家和建筑学家克里斯托弗 • 雷恩（Christopher Wren，1632—1723 年）爵士一起讨论了行星运行轨道的本质，以及能否用平方反比定律解释它。哈雷说他试过用该定律去计算行星的轨道，不过没有成功。胡克自夸说他计算出来了，不过拒绝向大家演示。雷恩有点不耐烦，同时心存怀疑。他于是向哈雷和胡克提出挑战：谁要是能在两个月之内给出演示，就奖给一本价值 40 先令的书。两个月的时间很快就过去了。当年 8 月，哈雷在剑桥大学再次见到牛顿时，与牛顿就该问题进行了讨论。这次讨论是牛顿一生中最具转折性的一个事件，由此也产生了西方科学和文化上最重要的事件之一——《自然哲学的数学原理》一书的出现。在这本书中，万有引力只是一个副产物。[13]

人类思想中影响最深远的一个归纳

古语说，法律就像香肠，最好别去看它们是怎么制作出来的——对

法律的来由知之越少，就会越发尊重法律给出的结果。与其说这是一句真理，还不如说它是一句妙言。你预期的结果是什么？如果能够真正理解人类的创造力，那么理解味美香肠的做法和法律的制定也就不是难事了。不过，这一名言倒是点出了创造力一个令人好奇的矛盾之处：重大事物的源头也可能是很卑劣的。在人类智力所取得的胜利中，很少有人能像牛顿得出万有引力定律的过程一样，将这个矛盾展现得如此犀利。这个过程中既有巨大的野心、口头的空喊和纠缠的秘密，也有积蓄的嫉妒心和赤裸裸的谎言。不过，结果却是惊人地奇妙。理查德·费曼曾说：“这是人类思想中影响最深远的一个归纳。”[14]

牛顿万有引力定律与第 2 章所讨论的力是在同一时期提出的。这项工作在牛顿还是剑桥三一学院的学生时就开始了。那时，牛顿常常在笔记本上匆匆记下自己对引力的各种评论。在有些评论中，牛顿把重力看成类似于原动力的能力。这种能力深居于物体内部，使物体发生运动。在其他一些处理天体运动的评论中，牛顿则融入了笛卡儿的解释：重力来自涡旋产生的粒子所造成的压力。在很长一段时间里，牛顿都接受离心力的思想—— 一个物体被另一个物体推开，就像拴在绳子上摆动的石头会受到绳子的拉力一样。

此后的 1680 年，因为两个重要事件，牛顿的引力思想发生了深刻的变化。这两个事件一个是哲学的，另一个是数学的。哲学思想从“认为力的原动力是使物体发生运动的内部推动力”，转变为“运动是因受到外部物体施加的力而产生的”这样一种观点。与此同时，人们开始逐渐认识到重量和质量之间存在着差异（其实在牛顿之前，波义耳、伽利略和开普勒已经在不同程度上认识到了这一点）。人们之所以会认识到这种不同，是因为力是随着距离而变化的。重力随着与地表距离大小的不同而发生变化，在不同的纬度上，物体的重量不同。然而，质量作为物体运动的关键元素，却是不变的。

另一个对牛顿思想产生深刻影响的重要事件是牛顿与他的对手胡克

的通信。二人的通信始于1679年。

牛顿讨厌胡克。1673年，胡克告诉皇家学会的同事，说牛顿最近关于光的开创性工作是错误的。胡克虽然言不符实，态度却很傲慢。这令牛顿颇为恼火，威胁说自己要彻底放弃科学研究。1679年秋天，胡克开始与牛顿通信，持续了两个月的时间。不过二人的通信也是同样地不妙。牛顿在第一封回信中犯了一个本不该犯的错误。于是胡克借题发挥，到处散播说牛顿误导了皇家学会的同事们。但是真正对牛顿构成挑战的却是胡克对平方反比定律和行星运动的质疑。胡克说行星沿着圆形轨道运行，并不是因为作用在行星上的离心力和向心力的共同作用，而是在于向心力和物体自身惯性的共同作用。这激发了牛顿的极大兴趣。

这种说法却使“牛顿走上了正轨”，尽管牛顿几乎倾其余生来否定胡克的贡献。[15]17世纪80年代早期，牛顿还没有提出万有引力定律。这一方面是因为他仍旧将彗星视为太阳系外的外来天体。不过他采用胡克的分析方法，将圆周运动分解为直线离心力和直线惯性效应，大获成功。从此，人们可以将所有物体（落体、行星）都理解为受到向心力的控制。牛顿还采用了胡克的方法和平方反比定律建立了开普勒各运动定律之间的基本联系。天体受另一天体的引力大小与两者的距离的平方成反比。天体沿着椭圆形轨道运行，中心天体则位于椭圆的一个焦点上，且中心天体与运行天体之间的连线在相同的时间内扫过相同的面积。[16]

1684年8月，哈雷路过剑桥并拜访了牛顿。与两人同时代的一个人是这么写的：

> 两个人呆了一会之后，（哈雷）博士问牛顿，如果太阳对行星的引力与行星离太阳距离的平方成反比，那么行星运动的曲线应该是怎样的？牛顿立即回答说曲线的形状应该是椭圆形。哈雷博士又惊又喜，问他何以知道事实如此。牛顿说：“为什么不是呢？我已经算过了。”于是哈雷博士要牛顿马上拿出结果给他看。牛顿在一堆论文中找了半天，没有找到。不过他答应重写一份，然后送给哈雷博士。[17]

这究竟是牛顿的妄想，还是秘密，亦或是他真的忘了计算结果放在哪里了？我们无从得知。不论牛顿到底是因为什么原因没能找到结果，从此他开始为哈雷重新计算。他于 12 月初整理出第一稿。稿子共 9 页，题目是《论运动》（*De Motu*）。牛顿在文中将太阳视为固定不变的。虽然太阳受到太阳系中所有其他物体的吸引，但却不受绕其旋转的行星的影响。这项工作将牛顿引领到了万有引力的门口，不过还缺少一个重要思想。根据牛顿第三运动定律——“有作用力就有反作用力，二者大小相等，方向相反”，如果太阳对行星有引力，那么行星对太阳也有引力。这样行星就会影响太阳的运动。这一点似乎是牛顿在完成了《论运动》的第一稿时才注意到的。

于是，牛顿开始对工作进行修正。这项工作于 1684 年 12 月底完成。修改稿首次对万有引力定律（所有物体对其他物体都有引力）这一重要思想进行了表述。表述采用了“所有这些行星彼此之间的相互作用”（eorum omnium actiones in se invicem）这一说法。如果某颗行星围绕太阳转动，那么二者将围绕一个公共的中心转动。但是太阳系中行星为数众多，每颗行星对太阳都有引力，且行星相互之间也存在引力。因此，行星运行的轨迹并不是完美的椭圆，两次运行的轨迹也不尽相同。牛顿曾这样写道：“的确，计算所有引力的最终结果太过复杂，超出了整个人类智力的能力范围。除非我错了。”[18]

牛顿不仅对太阳系有了更深刻的理解，而且也改变了科学程序。他把伽利略无限大平面上的无阻力理想实验推广到整个世界中。质量位于平面上，受力的作用而发生运动。科学家在此世界中创建模型（如开普勒运动定律），并将这些模型与实际观测结果相比较。但这些模型只是近似成立，尚需不断进行修正。牛顿早期的工作受到开普勒定律的启发。他当时认为开普勒定律是准确无误的。现在，牛顿通过这些模型指出开普勒定律是错误的，并预测出了定律与模型之间的偏差。[19]

1684 年 12 月，牛顿将《论运动》交给了哈雷。哈雷问牛顿能否发

表，牛顿没有同意。后来，牛顿又开始对《论运动》加以扩充，将新思想与其他思想（包括胡克在分析圆周运动时将其分为两部分的做法）一并融入到太阳系的结构中去。

1686 年，经过了 18 个月的工作，《自然哲学的数学原理》一书成形。该书是科学史上最重要的一本著作。在第一册的开头，牛顿灵活运用了胡克将圆周运动分为离心力和惯性的做法，推导出了开普勒定律和其他定律。在第二册的一部分中，牛顿表明笛卡儿的旋涡并不能解释行星的运动，并给出了一个恰当的解释。在第三册中，牛顿从上述解释出发，提出了“世界体系”。牛顿还进行了“月球测试”，测定出了地表上的物体所受的引力，并表明引力的强度与地球对月球引力的强度大小相同。而且，引力的强度与太阳和行星、其他行星和卫星之间的引力的强度也是相同的。牛顿在后续段落中明确提出了这些力。但直到今天，我们才把这些力称为“离心”力。这些力都是同一种力，人们赋予它们一个共同的名字：引力。引力“普遍存在于所有物体之中”。两个物体之间的引力大小与两物体的质量有关，“与两物体中心间的距离成反比”。人们现在把这一关系写成 $F_g = Gm_1m_2 / r^2$ 的形式。

胡克自称是发现引力定律的第一人。这么说也不是没有道理的，但是我们也知道为什么牛顿（和众多历史学家）驳回了胡克的请求。牛顿受益于胡克的工作这一点是很明显的。但是，牛顿说自己之所以看得更远，是因为站在巨人的肩膀上的时候，却也成了一个颇具讽刺意味的事实。牛顿暗示、讽刺胡克个子矮小，意即胡克所起到的推动作用更像是板凳，而不是高塔。胡克所提出的平方反比定律主要是针对一个物体的情形，或者至多是针对天体的情形。而牛顿却明确提出平方反比定律是普遍适用的。胡克对牛顿所起到的最重要的推动作用是使他知道了如何分析曲线圆周运动。然而，牛顿为了证明自己才是真正的发现者，在追溯万有引力发现过程的回忆录和对话集中做了假（包括月球测试）。这使万有引力定律的发现归属变得扑朔迷离。但是万有引力定律的发现者之

所以是牛顿，而不是布里奥、胡克或者其他人，乃是由于是牛顿首次明确提出了：引力不仅仅存在于某些相互吸引的物体之间，或者某些物体与吸引这些物体的其他物体之间；也不仅仅存在于落体与吸引落体的物体之间，或者是天体之间；而是普遍存在于所有物体之间。

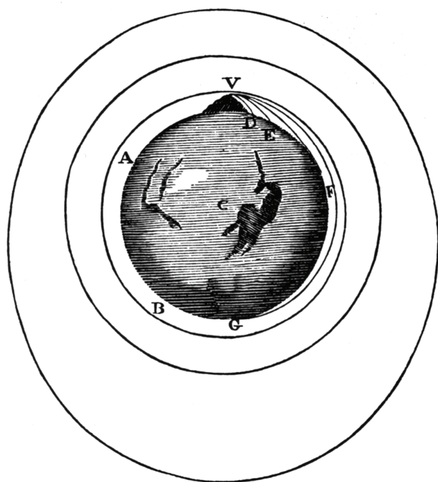
牛顿的解释也存在一些异常之处。例如，为什么受引力作用的物体的质量与公式 $F=ma$ 描述的推-拉力中物体的质量是相同的？真实情形可能未必如此。这是巧合吗？如果是巧合，那这巧合也真是够奇怪的了。这一问题的答案在 200 年后广义相对论的发展过程中发挥了作用。但是在牛顿时代，人们得仔细思考才能看出来。由此可见，牛顿的视野真是大胆新奇！

这种视野从深层次上来说是“一视同仁的”。万有引力是一种普遍存在的力，与物体的形状和物体在宇宙中的位置无关，而只与物体所含质量的大小有关。伽利略将所有的吊灯都看作是钟摆，从而获得了对事物的深刻理解。现在，牛顿更加雄心勃勃，他将所有物体都一般化，把它们看成是吸引子。对所有物体来说，重力都是随时随地存在着的。

解释定律的定律

人们把牛顿的万有引力方程看作西方科学中一个最重要变革的顶峰。牛顿本人由此也成为“黄金标准”。其他科学领域的学者常常将本领域的杰出人物与他比较。詹姆斯·克拉克·麦克斯韦（James Clerk Maxwell）把安培尊为电学领域里的牛顿；阿尔弗莱德·华莱士（Alfred R. Wallace）和托马斯·赫胥黎（Thomas Huxley）等人把达尔文看作是生物学领域的牛顿。

而且，人们常常引用牛顿的万有引力定律来说明成熟的科学所应具备的条件。弗朗索瓦·马让迪（Francois Magendie）在他的经典教科书《生理学速写》（*Elementary Sketch of Physiology*，1817 年）中遗憾地说，在他的领域中还没有出现“一个一流的头脑，能像牛顿发现引力定律那样发现生命力的定律”。



牛顿的炮弹理想实验。该图描绘了轨道的思想。如果在大气层之外的山峰上发射一颗炮弹，将会发什么现象？发射炮弹时对炮弹的推力越大，炮弹飞行的距离就越远。如果推力足够大，炮弹就会再次回到山峰上，并沿着相同的路径不断飞行下去。

然而，牛顿万有引力定律的影响已经超越了科学的范畴，扩展到了教育学、哲学、神学和其他人类文化领域。这一定律也使人们改变了对于“定律”本身的认识。

在现代，科学定律的概念有一个特定的含义：定律描述的是自然界和自然界的行为。例如，罗马城第三大学的哲学家莫罗·多拉拖（Mauro Dorato）在他的《宇宙的软件：自然定律的历史和哲学简介》（*The Software of the Universe: An Introduction to the History and Philosophy of Laws of Nature*）一书中将科学定律称为“物理系统的属性之间的数学联系”。

不过情况并非总是如此。对古希腊人来说，法律不是描述性的，而是规范性的（来自希腊语 nomos），即法律是人类的习俗或行为。法律是统治者给被统治者的命令，后者可以选择遵守或者不遵守。（对于世界中

的非人类部分，人们以定律形式理解的是事物特征本质的体现。）甚至直到 17 世纪，许多科学家还是拒绝将“定律”一词用到自然界中的常规现象上，并坚持认为这只不过是社会语言向自然界的形而上的延伸。但是随着人们逐渐认识到宇宙结构的规律性，笛卡儿等科学家开始倾向于将创世描述为超级立法者的立法行为。人类命令和非人类命令的不同之处在于，后者对于上帝的服从是无意识的，而前者的服从（或不服从）则是有意识的。

牛顿就是以上述方式看待世界的。他把自己看成是描述宇宙原理的人。宇宙原理遍布整个宇宙，影响着宇宙中的万物。它的影响是直接的、即时的和命令式的。这一原理的普遍适用性，以及牛顿在陈述“引力并不是物质属性”时所采取的谨慎态度，成为他如下观点的重要部分：他所描述的是超级立法者的行为。[20]牛顿的物质是无生命的，它只在受到力的作用时才发生运动。这样就“绕开了上帝”，保证了造物者不受任何约束。[21]牛顿的机械自然观（即物体被动地对外力做出响应）不仅与超级立法者相一致，而且也需要有这样的一位超级立法者。怎么可能只有法律而没有立法者呢？他曾这样写道：“最完美的行星和彗星体系无法人工制造，它只能由圣贤的强主发明和造出。”（牛顿可能认为这与太阳系的起源有关。太阳系起源现在可以很容易地由简单原理随时间的运动来解释。牛顿的想法让我们对要小聪明者的极端傲慢感到惊异，他们过于自信，认为如果无法解释事物的来源，那么就可以归结为上帝的作为。）

牛顿的万有引力方程极大推动了人们以描述性的视角，而非以规范性的视角看待定律的倾向。自然和社会两者之间的影响颠倒了过来：自然语言现在拓展到了社会领域。

牛顿的一个助手、皇家学会会员约翰·西奥菲勒斯·德萨居利耶（John Theophilus Desagulier）作了一首诗，题目是“牛顿的世界体系，最佳的政府模型”。德萨居利耶发现牛顿的世界体系中包括“最规律的普遍引力。它从太阳传播到所有行星和彗星的中心”，是（英国）政府“系

统的生动图景”，即通过有限的专制统治，就能构成对人们的自由、权利和基本人权的良好保障。德萨居利耶得出结论说，得益于这一点，“我们才能在君主统治下享有幸福”。这表明“引力在政治领域与在哲学领域中一样具有普遍意义”。[22]

然而，政治理论家也开始使用牛顿的语言，这极大影响了现代的民主概念。科恩在他 1995 年出版的《科学与科学之父：托马斯·杰弗逊、本杰明·富兰克林、约翰·亚当斯和詹姆斯·麦迪逊的政治思想中的科学》（*Science and the Founding Fathers: Science in the Political Thought of Thomas Jefferson, Benjamin Franklin, John Adams, and James Madison*）中详细说明了这一点。他指出，美国的建国者都读过牛顿，但杰弗逊无疑是唯一读过牛顿的《自然哲学的数学原理》的美国总统。杰弗逊的图书馆中有该书的好几个副本，并在墙上挂着牛顿的肖像。富兰克林年轻时受牛顿的影响很深，一直想在伦敦见一下牛顿。亚当斯曾在一次政治辩论中引用过牛顿的运动定律。而麦迪逊写过一篇短文，比较了自然界与人类社会的不同。

甚至连社会主义的诞生也与牛顿定律联系在一起。政治思想家亨利·圣西门（Henri De Saint-Simon，1760—1825 年）认为，牛顿定律不仅是最纯粹的科学思维范例，并且也为以博爱和联合组织为基础的人类社会科学的建立提供了一个模型。圣西门是社会主义的创立者之一。有一次，他梦见上帝向他指示坐在上帝手里的牛顿，并宣布世界应该由一个称为牛顿委员会的机构来掌管。除了改善人性之外，圣西门援引上帝的话说，该委员会的主要任务就是寻找“适用于社会体的新万有引力定律”。牛顿方程不仅是关键事实，而且是唯一的关键事实。它把科学统一起来，激发了人们去寻找既适用于个人和集体，也适用于国家之间的社会秩序定律。圣西门甚至责备牛顿未能把万有引力变成包罗万象的哲学体系。[23]人类发现社会秩序定律越早，并越早以此重新组织社会，就能越早获得解放。

诚然，圣西门是个浮夸的人物，属于那种妄自尊大的贵族—— 一个理想主义者、蹩脚的作家、傻里傻气的家伙，他还有些自我——这也是19世纪早期社会主义者都有的毛病。然而，这样的人并不止圣西门一个。其他一些政治思想家，包括皮埃尔·卡巴尼斯（Pierre Cabanis，1957—1808年）、查尔斯·傅里叶（Charles Fourier，1772—1837年）和乔万尼·莫雷利（Giovanni Morelli，1816—1891年）也尝试过把万有引力的思想应用于人类生活。他们坚信自由、自觉、能思考的个体是由普遍的、决定性的科学定律联合在一起的。卡尔·马克思（Karl Marx，1818—1883年）也曾受这一思想影响。

牛顿的万有引力方程的意义不仅在于它指出了物体（无论是卵石、宇宙飞船还是行星）之间引力的定量关系，还在于它启发了包括政治理论领域在内的其他领域的学者，去寻找描述性的普遍数学定律。如果说毕达哥拉斯定理告诉人们什么是证明，那么牛顿万有引力定律则告诉人们什么是定律。这样一来，万有引力定律不仅改变了人们对自然的理解，也改变了人们对社会生活和人类生活概念的理解。

万有引力方程至今依然是人类知识和理性的标志。在乔治·奥威尔（George Orwell）的小说《1984》中，主要人物温斯顿·史密斯（Winston Smith）向思想警察投降的最终标志是他否定了万有引力定律。史密斯承认 2+2=5，思想被完全击垮，停止了思考。

 茶 歇

牛顿的苹果

天上享用美酒佳酿的众神们，
与我一起向缪斯女神歌唱牛顿吧，
他找到了藏在真理中的宝藏。

福玻斯把神的光辉授赋予牛顿，令他拥有无尽智慧。
茫茫人世间，他是与众神最亲近的人。

——埃德蒙·哈雷（Edmond Halley），《牛顿颂》（*Ode to Newton*）

一个苹果能引申出什么呢？

牛顿看到苹果落地而发现万有引力定律的故事是最古老和人们最熟悉的科学传奇之一。[24]据说这件事发生于 1665 年或者 1666 年末，位于林肯郡索普村牛顿母亲的果园。此时牛顿为躲避剑桥大学的瘟疫而休学在家。长久以来，因为几个原因，人们一直把这个故事看成是虚构的。第一，这个故事像是做戏，不可能是真的。第二，牛顿的一位性格古怪却颇具影响力的早期传记作家大卫·布同达（David Brewster）对故事提出了质疑。第三，也是最重要的，故事所描写的不只是伟大的变革是如何发生的。故事中所隐含的因果关系一定是假的，因为牛顿看到苹果落地后，就在头脑中形成了万有引力定律，而没有什么进一步的分析和考虑。牛顿的自传作家理查德·韦斯特福尔（Richard Westfall）认为："这个把万有引力定律看作是灵机一动的故事把定律变得庸俗了。"[25]

不过，自传作家们已经发现了充足的证据。这些证据表明这个故事的最终源头是艾萨克爵士自己。他把这个故事告诉了几个不同的人，包括他的侄子（他又告诉了伏尔泰）和他的朋友威廉·斯蒂克利（William Stukeley，1687—1765 年）。以下是选自斯蒂克利回忆录的版本。

> 天气很暖（肯辛顿，英格兰，1726 年 4 月 15 日），我们来到花园的苹果树树荫下喝茶，只有我和他（牛顿）。我们聊着聊着，牛顿告诉我之前自己想到万有引力时的状况与现在完全一样——他正在沉思之际，一颗苹果从树上落了下来。他问自己，为什么苹果总是笔直地落到地上？为什么苹果不斜着下落或者向上运动，而总是向着地球的中心下落呢？[26]

但我们仍需保持怀疑的态度。为什么这个出了名地害羞、守口如瓶的人会突然变得健谈、滔滔不绝，把自己作出伟大发现的历程细细数来呢？这不像是牛顿的作风。其实，许多作家和历史学家都怀疑并非是牛顿狡猾阴险，要攻击胡克。胡克先是声称自己第一个发现了万有引力的平方反比定律，甚至还写了一封信给牛顿，要牛顿认可他是第一位发现者。牛顿通过这样一个故事，说明自己发现万有引力是 17 世纪 60 年代以前的事，于是胡克的说法自然就站不住脚了。这样的一个“骗局”更像是牛顿的风格，哪怕事实并非如此。

关于真正的情况，我们只能从牛顿某次被问及如何作出万有引力定律等发现时的回答中略知一二——“我总是想着它们，”牛顿说道，“我不断地去思考问题，直到黎明一点点展开，最后整个天空一片光明。”[27]这个说法听起来更合情合理，与除苹果故事之外的其他重大发现的产生过程也更加相似。“黎明”所包含的不光是能清楚看到的事物，还有人们从前人那里继承来的概念的变化和出现的新事物和新概念。伯纳德·科恩写道，“作为成功科学范例的万有引力定律的发现并非只是天才之笔”，而是包含了“现有思想的转变”的一个漫长过程。科恩还补充说：“万有引力定律的发现道出了我所认为的科学上的重大突破的一个基本特征，从最简单的创新到最重大的革命，那就是，通过变换已有思想，创造新的事物。”[28]

苹果在牛顿思考万有引力的过程中可能确实发挥了作用。如果事实的确如此，它所起到的作用就与苏格拉底向美诺的奴隶指出对角线一样是相同的。苏格拉底使奴隶小孩对这个他已有一定了解的问题有了新的理解，并在该过程中帮助他进行了思维的转化。

注　　释

[1] 另一个有趣的思想来自数学家帕普斯（公元前 3 世纪）。他提出了一种把引力像亚里士多德的“拉力”一样加以处理的方式。他说，先弄清要使平面

上的重物发生运动需要多大的拉力，然后倾斜平面，再弄清要使重物向上运动需要多大的拉力。Jammer，*Concepts of Force*，第 41 页。

[2] 林德伯格，《西方科学的起源》（*The Beginnings of Western Science*），第 275 页。

[3] 参见 Lynn Thorndike，《占星学在科学史上的真正地位》（*The True Place of Astrology in the History of Science*），ISIS 46 (1955)，第 273 页。

[4] Nicoletto Vernias，*De gravibus et levibus*，威尼斯 1504，引自詹摩尔，*Concepts of Force*，第 67 页。

[5] 哥白尼写道："我认为重力只不过是自然的某种嗜好。它是宇宙之神施于地球上的物体的神圣旨意，使他们能重新变成球形，回归统一和完整。可以想见，这种引力在太阳、月球和其他运行的天体上也是存在的。通过引力作用，天体就能保持人们所看到的球形。"《天体运行论》，第 1 卷，第 9 章。

[6] 行星沿着椭圆形轨道运行的思想是用已有基本原理获得的重大突破。开普勒对第谷的数据和哥白尼的日心说深信不疑。他认为第谷的数据大于理论值这一现象具有非常重要的意义，并对从古至今一直沿用的假设提出了质疑。这种信任使得实验数据与理论之间的偏差有了意义，同时让开普勒怀疑得恰到好处。在这里，信任在重大科学的发现中发挥了核心作用。不过这并不是第一次，也不会是最后一次。

[7] E. A. Burtt，《现代科学的形而上学基础》（*The Metaphysical Foundations of Modern Science*，纽约加登城：双日出版社，1954 年），第 64 页。

[8] 开普勒，《宇宙的奥秘》（*Mysterium Cosmographicum*），E. J. Aiton 译（诺沃克，康涅狄格：Abaris Books，1999 年），第 203 页。

[9] 1666 年，意大利生理学家 Alfonso Borelli（1608—1679 年）提出了一个木星卫星运动的解释。解释涉及了各种力之间的以及应用于行星运动的相互作用，指出这一定律对行星、行星的卫星、太阳和太阳系都是适用的。

[10] Carl B. Boyer，《科学家传记词典》（*Dictionary of Scientific Biography*）中的"伯努利，伊斯迈尔"词条，第 2 卷（纽约：Scribner's，1970 年），第 348~349 页。

[11] Robert Hooke，"Lectiones Cutlerianae"，见于 R. T. Gunther，《牛津早期的科学》（*Early Science in Oxford*），第 8 卷，1908 年，第 27~28 页。

[12] 胡克致牛顿，1679 年 11 月 24 日。牛顿，《牛顿信件》（*Correspondence of Isaac Newton*）第 2 卷，1676—1687 年，H. W. Turnbull 编（剑桥：剑桥大学出版社，1960 年），第 297 页。

[13] 牛顿的传记作家 Westfall 写道："科学史上没有几个阶段能取得像 1684~1685 年秋冬的这三到六个月之间那样的重大成就。"R. S. Westfall，

《永不停息：牛顿自传》(*Never at Rest：A Biography of Isaac Newton*，纽约：剑桥大学出版社，1988 年)，第 420 页。

[14] 费曼，《物理学讲义》，录音带 13，No. 1，第 1 面。

[15] 伯纳德·科恩，《科学美国人》(Scientific American)，1981 年 3 月。

[16] "是牛顿把开普勒的面积定律提升到了今天的地位。" 伯纳德·科恩，《科学美国人》，1981 年 3 月，第 169 页。

[17] 摘自伯纳德·科恩，《新物理学的诞生》(*Birth of a New Physics*，纽约：诺顿出版公司，1985 年)，第 151 页。

[18] 同上，第 236 页。

[19] 这是一个值得注意的发展，不过此种发展的模式常常在历史中上演：牛顿早期的工作受到了开普勒定律的启发；他认为开普勒定律准确描述了自然，可以令自己获得更深刻的洞察力。然而，牛顿获得的洞察力却表明开普勒定律是错误的。一来二去，牛顿预测出了开普勒定律的偏差。这一发展表明了人类智慧在科学上是如何依靠自己的力量，反复在"两个世界"中来回作用获得成功的；也表明了自然的外表如何，人们与自然打交道所用的概念是什么；以及这一过程是如何改变自然的外表和人们所用的概念的。"两个世界"指的是人们对世界的经验和世界模型。哲学家把这一过程叫作解释学。解释学是个专有名词，不过它表达的仅仅是基本的科学程序。这一过程常常会被隐藏起来，因为人们常常更容易注意到本质，而不是过程。可是如果没有这一过程，科学就会无足轻重或者成为不可能。

[20] 正如他在给 Bentley 的信中写道的："引力一定是由一种依照某种定律持续作用的媒介产生的。不过该媒介到底是物质的还是非物质的，还是留给读者去考虑吧。"牛顿，《牛顿信件》(*Correspondence of Isaac Newton*) 第 3 卷，1688—1684 年，H. W. Turnbull 编(剑桥：剑桥大学出版社，1961 年)，第 254 页。

[21] 历史学家 Marjorie Nicolson 对此感到困惑，是不是"牛顿觉得他提出的引力定律并不像古物那样，而是全新的呢"？她接着说："这里有一个最终证明，即微观世界能够反映宏观世界——控制宏观世界的行星和恒星、使其各行其道的定律与控制微观世界中比萨塔上落下的重物或者飞鸟翅膀上落下的羽毛的定律是完全相同的。在微观世界中，人处于核心位置。长久以来，古典的、中世纪的和文艺复兴时期的科学家、诗人和神秘主义者一直都在猜测整体和部分之间的这种重复、相互联系和互锁的关系。"Marjorie Hope Nicolson，*The Breaking of the Circle: Studies in the Effect of the "New Science"*

Upon Seventeenth Century Poetry（纽约：哥伦比亚大学出版社，1960 年），第 155 页。

[22] Westminster，1728 年。伯纳德 • 科恩在《科学和科学之父：杰弗逊、富兰克林、亚当斯和麦迪逊的政治思想中的科学》（*Science and the Founding Fathers. Science in the Political Thought of Jefferson, Franklin, Adams, and Madison，1955*）一书中讨论了这首诗，第 285~287 页。

[23] H. Saint-Simon，《圣西门选集》（*Henri Saint-Simon:Selected Writings*），K. Taylor 编（伦敦：Croom Helm，1975 年），第 78~79 页。

[24] 参见 D. McKie 和 G. R. de Beer，"Newton's Apple，"《伦敦皇家学会札记与纪事》（*Notes and Records of the Royal Society of London*）9（1951 年），第 46~54 页。

[25] Westfall，《永不停息：牛顿自传》（*Never at Rest：A Biography of Isaac Newton*，纽约：剑桥大学出版社，1988 年），第 155 页。

[26] William Stuckley，《艾萨克 • 牛顿爵士生平回忆录》（*Memoirs of Sir Isaac Newton's Life*，伦敦：Taylor and Francis，1936 年），第 19~20 页。

[27] E. N. da C. Andrade，《艾萨克 • 牛顿爵士的生活与工作》（*Sir Isaac Newton，His Life and Work*，纽约：Doubleday Anchor，1950 年），第 35 页。

[28] 伯纳德 • 科恩，"Newton's Discovery of Gravity"，《科学美国人》（*Scientific American*），1981 年 3 月，第 167 页。

4

数学之美的黄金标准
欧拉公式

$$e^{i\pi}+1=0$$

说明：自然对数的底（e）的 iπ 次方再加上 1 等于整数 0。其中 e 和 π 都是无理数；i 是–1 的平方根，是一个虚数。
发现者：伦纳德·欧拉。
发现时间：18 世纪 40 年代。

欧拉公式就像把握了爱情真谛的莎士比亚十四行诗，抑或是表现出人体内在之美的绘画一样，直指事物的最深处。

——基斯·德福林（Keith Devlin）

费曼 14 岁的时候第一次邂逅 $e^{i\pi}+1=0$。这位后来摘取了诺贝尔物理学奖的年轻人在日记中用大大的黑体字母写道："这是数学中最不简单的一个公式。"斯坦福大学数学教授凯思·德福林（Keith Devlin）这样写道："欧拉公式就好比是数学上的达芬奇的蒙娜丽莎画像或者米开朗基罗的大卫雕塑。"电气工程教授保罗·纳欣（Paul J. Nahin）在《欧拉博士的著名公式》（*Dr. Euler's Fabulous Formula*）一书中写道：欧拉公式为"数学之美确立了黄金标准"。与我保持通信的人当中有一人说该方程"美得如梦似幻"，另一人则把它称为"上帝的方程"。

18 世纪瑞士数学家欧拉所发现的这个公式已经成为了一个标志，它

具有特殊的性质。对于许多人，哪怕是只接受过有限数学训练的人来说，它已经超越了其本身所代表的事实。与许多标志一样，欧拉公式魅力无限、令人着迷。

欧拉公式是为数不多的在罪犯审判中作为证据出现的数学表达式。2003 年 8 月，一个环保恐怖主义者袭击了洛杉矶的多个汽车代理商，造成的损失价值 230 万美元。罪犯烧毁了一栋建筑，损毁和污损了 100 多辆 SUV 汽车。罪犯还涂写了一些标语，如“GAS GUZZLER”（油老虎）和“KILLER”（杀手）等。在一辆三菱蒙代罗汽车上，罪犯涂写上了公式“$e^{i\pi}+1=0$”。FBI 警方以此为线索，逮捕了威廉·克特雷尔（William Cottrell）。该线索后来也成为定罪的证据。克特雷尔是加州理工学院的理论物理专业研究生，他共受到纵火罪和阴谋纵火罪等 8 项指控。在 2004 年 11 月的法庭审判上，克特雷尔承认自己在那辆蒙代罗车上涂写了欧拉公式。此次庭审使克特雷尔获罪。在庭审过程中，克特雷尔说：“我 5 岁的时候就知道了欧拉定理。每个人都应该知道它。”[1]

另一个由方程演化而来的标志就是 $E=mc^2$。这个方程显然比欧拉公式更加为人所知。$E=mc^2$ 是人们所熟悉的流行文化的一部分，人们甚至把它建成了一座纪念碑。2006 年世界杯上，为体现德国是“思想之邦”，柏林建造了六座大型户外雕塑，其中有一辆汽车、一双球鞋，以及一个巨大的 $E=mc^2$。

可是，一个方程何以能成为一个标志呢？毕竟，公式还只是科学探索的第一步。欧拉公式无非是他在探索和研究方程的过程中得到的一个结论，而 $E=mc^2$ 是爱因斯坦在发展了相对论之后反复思考提出的。方程是否只是科学工具，与用方程来解决的任务相比，并没有什么内在价值？为什么有些方程能超越原本所属的科学探索过程，而获得内在的价值和重要意义呢？工具当然可以成为标志，好比斧头和镰刀就是前苏联的标志一样。但是像方程这样的数学和技术也可以吗？这样一个抽象的事物为什么能与一双球鞋和汽车并肩而立呢？

欧拉公式的故事将帮助对这些问题作出回答。

数学的分支

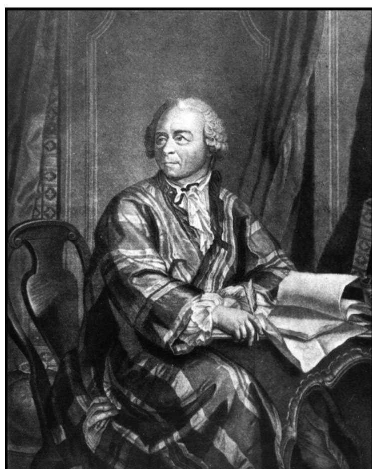
伦纳德·欧拉（Leonard Euler，1707—1783 年）是有史以来最多产的数学家。欧拉选集全部加起来将达 75 卷。他计算毫不费力，就像“人类呼吸、雄鹰翱翔一样自然”。[2]他有着惊人的记忆力，广泛涉猎各种知识，能一字不落地记下数学用表和维吉尔的《埃涅伊德》（*Aeneid*）整本。他能看到看似完全不同的数学领域之间的深层联系，并把这些联系表达出来，使结果看上去与 2+2=4 一样地自然。欧拉的基本公式都非常简洁优美。有一位注释者评论道：“方程的形式令人神清目爽。”[3]他的著名公式 $e^{i\pi}+1=0$ 就是最为简洁、完美，最令人清爽的一个。

欧拉生于瑞士的巴塞尔，父亲是一位牧师。父亲从小就教他简单的数学，激发了欧拉对数学的兴趣。欧拉上高中时仍有私人教师教他数学，因为学校里面不教这一科目。14 岁时，欧拉进入巴塞尔大学，学习神学、语言学和医学等，涉及面比较广泛。不过最令他着迷的还是数学。每周六下午，他的私人教师、著名数学家约翰·伯努利（Johann Bernoulli）教他数学。欧拉后来还与伯努利的儿子尼古拉斯和丹尼尔成为了朋友。1723 年，欧拉拿到学位后，遵从父亲的意愿，打算做一名神学家。不过很快他又开始研究数学。

研究数学并非易事。当时，大学里的研究基本都是在人文领域，留给数学家或者科学家的位置很少。而这些有限的位置也常常为皇家学院所掌握。

幸运的是，俄国彼得大帝和他的第二任妻子凯瑟琳一世（历史上有名的“文艺复兴夫妻”）正在圣彼得堡着手建立俄国科学院，并在整个欧洲范围内招募著名科学家。早先招募到的是尼古拉斯和丹尼尔。这两人后来又邀请了他们的朋友欧拉。1727 年欧拉到达俄国科学院，但此时彼得大帝和凯瑟琳都已去世，他们的继承人对于科学院的热情并不高，不

过欧拉仍受到了照顾和支持。周围都是一流的科学家，而欧拉很快成为了科学院的首席数学家。欧拉极为多产，科学院期刊的编辑把他的手稿堆成堆，桌上一有空间就从上面取一些下来。在俄国科学院的 14 年间他遭受了一些苦难，这其中最大的就是右眼失明。这可能是由过度工作导致眼睛疲劳而造成的。然而，在这段日子里，他自由地进行了大量的计算，在该过程中重塑了数学的基础。



伦纳德 • 欧拉（Leonard Euler，1707—1783 年）

数学发展的方式常常不是直接的，就像城市的发展一样。人们先是建起一些定居点，它们之间几乎没有影响。后来这些定居点聚集在一起，形成住宅区。但住宅区的形成是随机的，适应性较差，几乎没有什么商业。此后出现了一位具有远见的领导者，他对各住宅区非常了解。通过重新命名街道，在重要的中心之间修建新的街道，形成了较大的、更加复杂、更有组织和更加统一的建筑。

欧拉在 18 世纪的数学中所扮演的角色就是上面这位具有远见卓识的领导者。

此时，数学中有两个发展成形的分支领域——代数和几何。**几何**研究的是点、线、面和由这三者构成的图形的性质。这些已经在古代欧几里得的《几何原本》（约公元前 300 年）中得到了系统的阐述。几何的一个分支是三角，它研究的是三角形的角度与边长之间的关系。三角最先是作为天文学的工具出现的。代数研究的是具有有限元素和离散解的方程。它主要研究有理数：能用整数或整数之比表示的数（p/q 的形式），亦即小数部分不断重复的数。（像 π 之类的数，小数点后的数有无限多位，而且不重复，称为无理数。）早在中世纪的时候，阿拉伯数学家穆罕默德·伊本·穆萨·花拉子米（Muhammad ibn Mūsā al-Khwārizmī，约 780—850 年）就对代数进行了很好的整理和组织，并为其定名。花拉子米对数学的贡献集中反映在《还原与对消的科学》（Hisāb al-jabr wa'l-muqābala ，830 年）一书中。他采用“还原”（al-jabr）一词来表示在方程的两端加上相等的量，对其进行简化的过程。“还原”直译为希腊语就是“代数”（algebra），自此“代数”也成为整个领域的代名词。

分支的统一

18 世纪初出现了一个新的数学分支，称为数学分析。数学分析研究的是包含无穷多项的无穷序列的分支，也可以说数学分析是研究无穷序列的技巧的集合。它在很大程度上是由微积分发展而来的。微积分研究的是连续过程，由戈特弗里德 • 莱布尼茨（Gottfried Leibniz）和牛顿（他称之为流数论）发明的[①]。数学分析还研究无理数和虚数，即负数的平方根。虚数由哲学家和数学家笛卡儿命名，似乎他认为虚数是虚构的。虚数后来在数学上的用途和价值逐渐增大，但这个名称却一直沿用下来。

但是，经过欧拉的组织，数学分析成为了完整的知识体系和有组织的新生数学领域。例如，欧拉首次对函数进行了系统的研究。函数现已

① 关于是牛顿还是莱布尼茨首先发现了微积分，是科学史上一桩著名的公案。
——编者注

成为不可缺少的数学工具，它把一个数与另一个数匹配起来。函数的简单例子有计算税率的公式和从华氏温度转换为摄氏温度的公式等。欧拉还提出并扩展了数学家用以计算无穷项级数之和的数学工具。在欧拉之前，数学家一直视计算无穷项级数的和为苦差。只有在没有其他方法可用时，才会采用此法解决问题。欧拉告诉数学家们，无需惧怕这样的级数。只要级数是收敛的，级数的和就很容易计算。欧拉还是数学史上最具影响力的数学符号提出者。他引进或标准化的重要符号有以下这些。

- π，圆的周长与直径之比，可能取自希腊语“周长”的第一个字母。
- e，自然对数的底，可能取自“指数”的第一个字母。对数就是底要经过多少次方才能变成一定值，e 是自然对数的底（$\log_e y = x$ 的意思是 $e^x = y$）。[4]
- i，即−1 的平方根，基本“虚数”。i 并不像笛卡儿想象的是虚构的。它拓展了可解方程的范围。[5]
- $f(x)$，x 的函数。函数将一组数与另一组数匹配起来。
- sin，正弦函数的简写，它把直角三角形中一个角的度量和该角对边的长度与斜边的长度之比匹配起来。
- cos，余弦函数的缩写，它把直角三角形中一个角的度量和该角邻边的长度与斜边的长度之比匹配起来。
- Σ，用于计算级数的和。

1741 年，在圣彼得堡呆了 14 年的欧拉应另一个文艺复兴式的人物——腓特烈大帝的邀请，来到柏林科学院。不过他与圣彼得堡的同事仍保持着密切的通信联系。欧拉发现柏林并不如圣彼得堡更适合自己。腓特烈大帝已经习惯了那些卖弄学问的人，认为沉默寡言、不善表现的欧拉是他所网罗到的博学者中的一个例外，称他是“数学领域的库克罗普斯①”。[6]1766 年，在柏林呆了 15 年之后，欧拉应凯瑟琳大帝的邀请，

① 库克罗普斯是希腊神话中的独眼巨人。——编者注

再次回到圣彼得堡。尽管在圣彼得堡得到了很大的支持，欧拉的健康状况还是不断恶化。他自己知道另一只眼睛的白内障正在不断发展，最终可能会失明。对此，他勇敢面对："如此一来，我就更不容易分神了。"他开始学着在石板上用粉笔书写，再让孩子们誊抄下来。这样一来，注意力确实会更加集中。接下来的 17 年间，欧拉在数学的天地里继续无畏前行，不断计算、修改、写作，绕着桌子边走边说。儿子和助手就记录下他所说的话。欧拉就是在这种完全失明的情况下，完成了他几乎所有工作的另一半。

1771 年，一场大火摧毁了大半个圣彼得堡，欧拉的房子也着了起来。此时的欧拉身体虚弱，而且双目失明。是朋友把他从房子里扛出来，送到了安全的地方。就在朋友扛着他向外跑的时候，欧拉依旧没有停止计算。1783 年 9 月 18 日，欧拉教小孙子数学时，计算出了热气球的路径和新近发现的行星天王星的可能轨道等问题。突然间，欧拉的烟斗从嘴里滑落到了地上，同时，"停止了计算和呼吸"。[7]

今天的数学大厦与欧拉时代的相比，要高大得多。如今，大的数学分支有分析、代数和拓扑学。在这三个领域中，欧拉都起到了推动作用。他写的数学教科书《完全代数学导论》[*Vollstandige Anleitung Zur Algebra*，在英国出版的版本是《代数学基础》(*Elements of Algebra*)]，对代数领域的介绍与现在的基本一致。他也是最先涉足拓扑学的学者之一，尽管当时还没有这门学科。他给出了哥尼斯堡七桥问题的一个著名的解。该问题说的是能否一次走过连接城市的河岸和小岛的七座桥，而且不能再次经过任何一座桥。大约 100 年之后，拓扑学被人们接受，成为数学领域的重要分支。

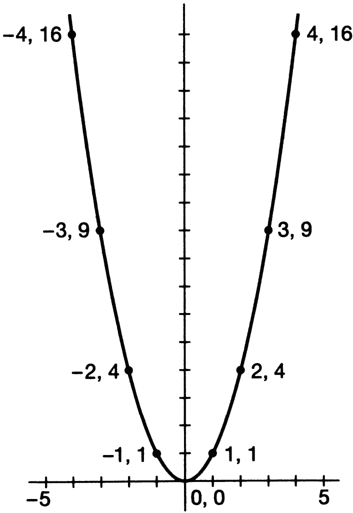
尽管如此，人们公认欧拉是数学分析大师：学者们常称他为"数学分析转世"。欧拉在该领域最重要的工作是在柏林期间撰写的一本两卷教科书，书名是《无穷分析导论》(*Introductio ad analysin infinitorum*，1748 年，以下简称《导论》)。在书中，欧拉给出了函数的许多发现（包

括无穷级数）以及之前未被证明或者证明不够完整的定理的证明。欧拉还提出了一些定义和符号，它们后来很快就成为了标准，包括 π 和 e。"《导论》一书对数学分析的重要性就好比是欧几里得的《几何原本》对几何的意义和 al-Khowârizmî 的 *Hisâb al-jabr wa'l muquâbalah* 对代数的意义。数代人受这部经典著作启发，开始学习数学分析，特别是无穷级数。"[8]

但是，《导论》的意义并不只是重新对数学分析加以组织。它把许多数学符号和数学公式变成了无穷级数的语言，使数学分析从一个全新的、处于不断发展中的领域，一跃成为与已有的几何学和代数学并行的主要数学领域。

深层联系

在《导论》一书中，欧拉宣布了一个重大发现，即指数函数、三角函数和虚数之间的深层联系。这一证明源自欧拉对指数函数的研究。简而言之，指数函数包括一个叫做"底"的数和另一个位于"底"的右上角的数，即指数。指数表示"底"要与自身相乘多少次才能得到函数值（这种记号由笛卡儿发明）。指数函数的简单形式是 $y=2^x$，其中 2 是底，x 是指数。对任意整数 x，都可由该函数得出有限的级数项和整数乘积。例如，$2^2=2\times2=4$，$2^3=2\times2\times2=8$，$2^4=2\times2\times2\times2=16$，等等。



这些整数对可以拟合到一条曲线当中。在一条有着无穷多个点的曲线上，只有有限个点是由整数对组成的；这些点之间的曲线是由像 3.81

这样的小数或者像 $\sqrt{2}$ 和 π 之类的无理数组成的。这样一来，令 2 乘以自身 2.31 次，$\sqrt{2}$ 次和 π 次是什么意思呢？有理数可以用 p/q 的形式来表达，因此 2 的有理数次方的意思就是 2 的 p 次方的 q 次平方根。例如，2 的 3.81（=381/100）次方就是 2 的 381 次方的 100 次方根。而某数的无理数次方就是填充在曲线上除有理数次方之外的那些点，可通过计算无穷序列的极限得到。所以，2 的 π 次方就可以由 $2^3, 2^{3.1}, 2^{3.14},\cdots, 2^{3.1415926}$ 计算得到，序列中所取的 π 小数点后的位数逐渐增加。

在《导论》一书的第 7 章，欧拉表明，如果用下列无穷序列的加和来作指数函数的底，在数学上将会有很多好处：

$$1+1+\frac{1}{2!}+\frac{1}{3!}+\frac{1}{4!}+\frac{1}{5!}+\frac{1}{6!}+\frac{1}{7!}+\frac{1}{8!}+\frac{1}{9!}+\frac{1}{10!}+\frac{1}{11!}+\frac{1}{12!}\cdots$$

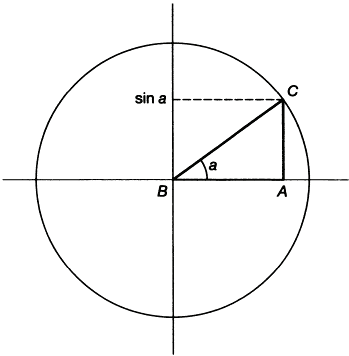
欧拉注意到，这些项的和是无理数 2.718281828459…。为简单起见，欧拉用 e 来表示这个数。e 是自然对数的底，也是最重要的数学常数之一。后来欧拉又注意到，如果把 e 作为自然对数的底，那么对于任意的 x，函数 e^x 可用以下无穷序列计算得出：

$$e^x=1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\frac{x^4}{4!}+\frac{x^5}{5!}+\frac{x^6}{6!}+\frac{x^7}{7!}+\frac{x^8}{8!}+\frac{x^9}{9!}+\frac{x^{10}}{10!}+\frac{x^{11}}{11!}+\frac{x^{12}}{12!}\cdots$$

上述序列称为指数函数，是所谓的泰勒级数的一个例子。[9]

在第 8 章中，欧拉开始研究三角函数。他从直径为 1 的圆的周长为无理数 3.141 592 65…这一基本事实出发。为简洁起见，欧拉将其称为 π。接着欧拉描述了三角函数的性质。三角函数涉及的是直角三角形中的角，用各边边长之比表示。例如，正弦函数涉及直角三角形中的锐角，以角的对边与斜边之比来表示。正弦函数可通过以下方法从锐角推广到任意角度：在（x,y）平面上画出一个直角三角形 ABC，使斜边 BC 的长度为 1，顶点 B 落在原点（0, 0）处，顶点 A 落在正 x 轴上，顶点 C 落在 x 轴上方。设 a 为 $\angle ABC$，测量时从正 x 轴沿着逆时针的方向进行。因此，$\sin a$ 即为 AC/BC。因为 BC=1，所以 $\sin a$=AC 的长度=点 C 的 y 坐标。如果

我们用“点 C 的 y 坐标”表示 $\sin a$（a 为 $\angle ABC$），那么就可以得到一个对任意角度都成立的定义：令 BC 转过一定的角度 a（从 x 轴正半轴出发，沿逆时针方向转动），并记下点 C 的 y 坐标。相应地，角度 a 的正弦值就会从 0 变到 1（90 度），再变到 0（180 度），再变到−1（270 度），最后再变到 0（360 度），并在之后的 360 度循环中重复以上变化。由此得到类似于示波器上所谓的“正弦波”的波形。余弦函数的定义相同，只是将点 C 的 x 坐标作为余弦值。随着角度的变化，余弦值从 1 变到 0，从 0 变到−1，从−1 变到 0，再从 0 变到 1，如此往复循环。余弦函数的图形与正弦函数相同，只是二者不同相。



之后，欧拉发现了正弦和余弦函数的几个比较明显的性质，包括通过简单应用毕达哥拉斯定理就能得出的一个事实，$(\sin x)^2+(\cos x)^2=1$。

欧拉在继续总结牛顿和其他前人工作的基础上，又进一步发现正弦和余弦等三角函数可以用无穷级数表示。例如，函数 $\sin x$ 可用如下无穷多项的和来表示（为简单明了起见，我们采用**黑体**来书写这些项）：

$$\sin x = x-\frac{x^3}{3!}+\frac{x^5}{5!}-\frac{x^7}{7!}+\frac{x^9}{9!}-\frac{x^{11}}{11!}\cdots$$

对于余弦部分，我们则用**加粗的黑体**来表示：

$$\cos x = \mathbf{1}-\frac{\mathbf{x^2}}{\mathbf{2!}}+\frac{\mathbf{x^4}}{\mathbf{4!}}-\frac{\mathbf{x^6}}{\mathbf{6!}}+\frac{\mathbf{x^8}}{\mathbf{8!}}-\frac{\mathbf{x^{10}}}{\mathbf{10!}}\cdots$$

采用这些函数，欧拉表明其他三角函数也可以类似地用无穷级数来表示。

欧拉精于计算，他可以把三角函数加和起来，得到以 e 为底的指数函数。在计算的过程中，欧拉采用了虚数 $\sqrt{-1}$ 。在《导论》写成的几年

之后，欧拉用 i 来表示 $\sqrt{-1}$ 。i 虽然不是“实”数，在数轴上也没有位置，但却常被用于实数运算。这样数学家就可以求解原本不可解的方程。如果将 i 插入到 e^x 的指数项中，则在无穷级数的所有项中都会出现 i：

$$e^{ix}=\mathbf{1}+ix+\frac{(i\boldsymbol{x})^2}{\mathbf{2!}}+\frac{(ix)^3}{3!}+\frac{(i\boldsymbol{x})^4}{\mathbf{4!}}+\frac{(ix)^5}{5!}+\frac{(i\boldsymbol{x})^6}{\mathbf{6!}}+\frac{(ix)^7}{7!}+\frac{(i\boldsymbol{x})^8}{\mathbf{8!}}\cdots$$

由于 i^2 等于−1，因此有 $i^3=-i$、$i^4=1$、$i^5=-i$，等等。于是上述级数可以变为：

$$e^{ix}=\mathbf{1}+ix-\frac{\boldsymbol{x}^2}{\mathbf{2!}}-\frac{ix^3}{3!}+\frac{\boldsymbol{x}^4}{\mathbf{4!}}+\frac{ix^5}{5!}-\frac{\boldsymbol{x}^6}{\mathbf{6!}}-\frac{ix^7}{7!}+\frac{\boldsymbol{x}^8}{\mathbf{8!}}\cdots$$

欧拉发现如果将 i 的多项式并为一组，就可以得到：

$$e^{ix}=\left(\mathbf{1}-\frac{\boldsymbol{x}^2}{\mathbf{2!}}+\frac{\boldsymbol{x}^4}{\mathbf{4!}}-\frac{\boldsymbol{x}^6}{\mathbf{6!}}+\cdots\right)+i\left(x-\frac{x^3}{3!}+\frac{x^5}{5!}-\frac{x^7}{7!}+\cdots\right)$$

或者，就像他在《导论》的第 8 章中所写的那样（用 i 表示 $\sqrt{-1}$ ，英文译本同此）[10]，为：

$$e^{ix}=\cos x+i\sin x$$

上述方程在指数函数和三角函数之间建立了深层次的联系。印度著名数学家拉马努金（Srinivasa Ramanujan，1887—1920 年）在上中学时曾发现了这个关系，并激动地把它写了下来。不过，拉马努金在得知自己并不是第一个发现这个关系的人之后，非常沮丧，把计算结果藏到了自己家房子的屋顶上。[11]

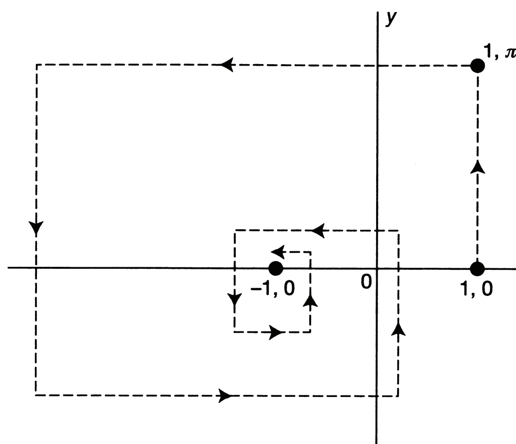
这个方程令人不可思议，不过它的意义还远不止于此。假定 x 等于 π，由于 $\sin\pi=0$，$\cos\pi=-1$，因此 $e^{i\pi}=-1$ 或者 $e^{i\pi}+1=0$。

还有一种方法是用图形的方式证明上述方程正确性。假定用 π 代替 x 代入前几段的公式中，则该公式变成：

$$e^{i\pi}=1+i\pi+\frac{(i\pi)^2}{2!}+\frac{(i\pi)^3}{3!}+\frac{(i\pi)^4}{4!}+\frac{(i\pi)^5}{5!}+\frac{(i\pi)^6}{6!}+\frac{(i\pi)^7}{7!}+\frac{(i\pi)^8}{8!}\cdots$$

数学家可以把上述各项像向量一样加起来。各向量首尾相接，带有

虚数 i 的项表示该向量逆时针转过 90 度。[12]如果从 0 出发，第一项（即 1）是沿着 x 轴出发到达点（1, 0）的向量。第二项（iπ）是一个从（1, 0）出发，相对于第一个向量逆时针转过 90 度，向上延伸 π 个单位，到达坐标（1, π）的向量。第三项（$i\pi^2/2!$）是从（1, π）出发，相对于第二个向量再逆时针旋转 90 度，方向与第一个向量完全相反，并越过直线 x=0 到达点（−（$\pi^2/2-1$）, π）的向量。第四项是一个方向向下的向量，延伸至 x 轴的下方，如此往复。由于后一向量总是沿着前一向量逆时针旋转 90 度的方向，并且分母增长的速度比分子要快，导致向量的模①逐渐变小，所以最后就能得到收敛于点（−1, 0）的多角形螺旋（见图）。



表明无穷级数收敛于−1 的多角形螺旋

这种简单的欧拉公式（根据某些定义，此种形式的欧拉公式并不是方程，因为它不含有变量）包含了 5 个数学上最基本的概念——0、1、自然对数的底 e、虚数 i 和 π，还包含了 4 个运算符——加、乘、取幂和相等。这些概念和运算符分别只出现了一次。欧拉公式说的是无理数的

① 向量的长度或者绝对值。——译者注

虚数乘无理数次方加上 1 恰好等于 0。数字 π^e、2^π 和 e^π 都被认为是无理数。但 $e^{i\pi}$ 却占据了数字大厦中的一个特殊位置。在这个位置上，无理数和虚数结合在一起，幽灵般地“相互抵消”，产生出了 0。曾有人说，所有的分析都集中在这一个公式当中。[13]欧拉的结果与其他一些事实表明：受到笛卡儿嘲笑的虚数在数学中并没有被边缘化，而是位于数学的中心。它们在数学上将发挥越来越重要的作用。伴随着 20 世纪量子力学的发展，在物理学和工程学等研究周期性现象的领域中，复数也被用于表示波。有了复数就能同时表示相位和波长两个过程，而复指数则可以将直线映射为复平面上的圆。

也许，欧拉公式 $e^{i\pi}+1=0$ 只代表了一种含义，只是方程探索之路上的一步。欧拉公式“只是”一个方程，是科学探究这一持续过程中上千步中的一小步，只是欧拉对方程进行广泛研究后得出的一个含义。但是，这几千步中的某几步会得到，而且也应该得到特殊的地位。在仍处于发展和变革中的重要科学领域内，有些表达式将起到里程碑式的作用。它保留前人的工作，着眼于现在，面向未来。理论、设备和人一直会变，可是方程和公式却是几乎不变的。它们指导人们如何解决问题，是人们设计新设备的工具。专家们则用它们描述新的发现。公式和方程把知识加以总结，保存下来，并期待和展示新的发现。

欧拉公式也体现出欧拉是如何对数学进行重筑的。与其他科学一样，数学并不是沿着预定的轨迹发展的。数学的发展具有历史上的偶然性，一代代的科学家从前人那里把假设、技巧和概念继承下来，加以改造，再传给后人。有了这个过程，人们才能以特定的方式有条理地去理解数学。就像有一个实体论，可以将不同的现象归结到完全不同的领域中去。所有方程都会隐含地采用这一继承下来的数学架构。不过，欧拉却对实体论重新加以组织，以分析为中心，两边是几何和代数。回首过去，数学家可能会把数学上最新的重组都看作是不证自明的。无疑，这也正是数学家卡尔·弗雷德里希·高斯（Carl Friedrich Gauss）说“谁要是觉得

$e^{i\pi}+1=0$ 不是显而易见的，谁就不是数学家”的原因所在。一个人如果无所不知，那么就没有什么能让他感到惊奇的了。但是，数学家也并非天生就是数学家。他们小的时候不是数学家，也得一点点地学。在这样的学习过程中，他们常常会对自己获得的并不全面的数学知识进行变换和重组。欧拉的简明公式 $e^{i\pi}+1=0$ 就是对这一过程最简洁的描绘。[14]

欧拉公式成为标志性的事物还有另外一个深层的原因。德福琳这样描写它：“欧拉公式深入到了实体的最深处。它把源自人们生活不同方面的思维抽象出来，合而为一，并再次提醒人们，联系、结合在一起的事物比相互分开的事物更重要、更有价值，也更加绚丽多姿。”

德福琳的评论表明，为什么欧拉公式能超越得出它的特定科学探究过程，变得更有价值和意义。这个例子很清楚明白，因为它说明了方程和公式的作用：它表明了看似不相干的，甚至是矛盾的元素（有理数、无理数和虚数）是如何包含在一个整体之中的。而且这一过程可以说没有几步运算，非常简洁。它一下子就把各个元素简化、组织和统一起来了。欧拉公式向人们展示出方程的作用，以及它何以能成为方程。



茶　歇

标志性的方程

记者：俄国人有 GISMO（通用矩阵解算装置）吗？
科学家（罗德·泰勒）：没有，不过我确信他们会喜欢 GISMO。
记者：您能给出该方程吗？
科学家：不能，我确信他们会更喜欢这个方程。

——影片《谍海玻璃船》
（*The Glass Bottom Boat*，1966 年）中的对白

方程对于我们的知识与思想结构的影响远远超出了科学的范畴。披着数学外衣的一些冠冕堂皇的思想常常会令其看上去更具权威性、更加确定、准确和持久。笑话、箴言、政治口号和催人奋进的励志标语也常常被改头换面,以方程的形式出现:“知识=力量”、“战争=屠杀”、“准备+耐心=成功”。方程写的是幽默,就像 2007 事件中所分析的一样。在该事件中,节目主持人丹·伊姆斯(Don Imus)发表了种族歧视的言论,被 CBS(哥伦比亚广播公司)解雇。他说:“白皮肤的人加黑色标语等于喜剧。不过这个方程也有失效的时候。白皮肤的人加黑色标语减常识等于悲剧。”[15]

或者考虑一下乔治·奥韦尔(George Orwell)的小说《1984》中的著名方程。虽然这些方程明显是错误的,不过它们却指出了另一类不同的事实:

战争=和平

无知=力量

自由=奴役

虽然上面的方程表面上与数学和科学有着相似之处,不过其实它们只是伪装的比喻而已。“=”号并不是数学上所说的“等于”或者是等价。在数学上,“=”是定量的,意思是“完全相等”。它指的是一个集合中所含项目的数量,或者是一个特定的可测的量。以上是数学的基础。举例来说,知识就是力量这种方式就有很大不同。它是定性的,只能通过引入看似不证自明的“相同”、“相等”和“是”等词的含义的哲学复杂性来表达。

这些奇异的方程也是很有趣的。它们体现出了一种危险的期望:其他各种知识也可以通过相同的项,用简洁的形式、相同的数量和简单的单位来表达。也就是说,方程诱导人们认为这是我们应该使用的思维方法,而其他的方法都是低级的或者是有缺陷的。某个科学杂志的一位记者收到了一家早餐麦片粥制作公司的一个人的一封电子邮件,请他提出

一个方程，确定何时添加牛奶是最佳的。这位记者对大众似乎过于着迷，甚至要为最琐碎的一些小事找出方程来的做法很不以为意。这封信也引出了其他人的一些评论。这些人见过有人要求提供制作三明治、车辆停放和完美情景喜剧的方程。评论警告说，这种做法有其负面影响，原因不光因为它们是伪科学，还在于这种做法鼓励了科学家不负责任的行为，也误导了公众对科学本质的认识。[16]

某些特定的方程可能具有很宽泛的象征意义。以 1+1=2 的小老哥 2+2=4 为例。在虚幻和现实中，2+2=4 被用于象征非理性之于理性的优越性、理性之于非理性的优越性，以及神之于非理性和理性的优越性。[17]如在陀思妥耶夫斯基（Dostoyevsky）的小说《地下室手记》（*Notes from Underground*）中，解说员将 2+2=4 描述成"难以忍受的"、"有点傲慢"、枯燥、理性、没有生气。解说员发现它比"二乘二等于四要高级得多"。另一方面，在小说《1984》中，主要人物温斯顿把 2+2=4 作为不证自明的事实与明智和理性的试金石。它在思考时可以随时用到，为抓住客观事实指明了道路，帮助温斯顿确保客观事实是存在的。对党派而言，2+2=4 是必须用思想矛盾和党派执政的成功击垮的最后一点阻力，标准以外的必须彻底除掉。奥韦尔只是引用了前苏联领导人的一句口号，即他们把 2+2=5 写在宣传栏和电灯上，作为乐观、人定胜天和"只凭力量的魔力就能创造奇迹"的象征。[18]似乎正确的方程反而是错误的，它枯燥、理性、陈腐，不能捕捉到人类的创造力；而错误的却是正确的，它象征着人类的创造力可以克服自然的局限。同时，建筑师和发明家巴克明斯特·富勒（Buckminster Fuller）喜欢用"1+1=4"这一格言定义协同效应。也就是如果能有效地、创造性地将部分组合起来，就能产生用传统办法无法获得的效果。最后，牛津著名的神学家玛丽莲·麦科德·亚当斯（Marilyn McCord Adams）认为"人被创造出来，并不是要单打独斗的，而是要与上帝建立无所不在的伙伴关系"。她说这个"谦卑的灵魂"是"创造性的洞察力的助产士，他巧妙地推动、启发、指引着人们，直到人们发

现 2+2=5”。[19]

用奇异方式利用方程的小说家有卡尔维诺（Italo Calvino），他在《宇宙连环图》（*Cosmicomics*）一书中把爱因斯坦的广义相对论融入了一个故事中。另一个作家是马克·林内（Mark Leyner），他的《还有你，宝贝？》（*Et Tu, Babe*）一书中有个角色自称在阳物上刺着普朗克的能量公式 $E=h\nu$，也就是与辐射和能量相关的公式。此人后来不得不当着法官的面承认纹身的那个公式其实是伽利略落体定律：$d=16t^2$。这令他深感羞愧。

如果说方程也有缺点，那就是它们容易使人们认为知识就存在于方程中，而不是存在于不断完善、更新的科学（柏拉图将此称为多提问）大厦中。方程会助长错误的观念，即科学是由一些需要记住的事实和信念组成的，而不是通过超越已有事实和信念，寻求对自然的更多理解的过程。

注　　释

[1] Ed Leibowitz，“The Accidental Ecoterrorist”，《洛杉矶》（*Los Angeles*）杂志，2005 年 5 月，第 100~105 页，第 198~201 页。

[2] 引自 Carl A. Boyer，《数学的历史》（*A History of Mathematics*，普林斯顿：普林斯顿大学出版社，1985 年），第 482 页。

[3] Marquis de Condorcet，“Eloge to Mr.Euler”，J. Claus 译，

[4] Martin Gardner，《意料之外的绞刑和其他数学娱乐》（*The Unexpected Hanging and Other Mathematical Diversions*，芝加哥：芝加哥大学出版社，1961 年）中的第 3 章有关于 e 的精彩论述。

[5] 费曼、雷顿和山德士的《费曼物理学讲义》（*The Feynman Lectures on Physics*，纽约：安德森·威斯利出版社，1963 年）第 1 卷有两节（22-5 和 22-6）关于虚数和虚指数的精彩论述。

[6] 引自 Boyer，《数学的历史》（*A History of Mathematics*），第 493 页。

[7] Condorcet，“Eloge to Mr.Euler”。

[8] David M. Burton，《数学的历史》（*The History of Mathematics*，纽约：麦格劳希尔出版公司，1985 年），第 503 页。

[9] 对于任意的 x，要得到 2^x 的值，可以将 2 乘以自然对数 ln(2)，然后以 $x\ln(2)$ 为指数：$2^x=e^{x\ln(2)}$。

[10] 欧拉，《无穷分析导论》（*Introduction to Analysis of the Infinite*），第 1 册，J. D. Blanton 译，纽约：施普林格出版社，1988 年，第 112 页。欧拉于 1743 年首次发表于 *Miscellanea Berolinensia 7*，第 179 页。

[11] G. H. Hardy，P. V. Seshu Aiyar 和 B. M. Wilson 编，《拉马努金论文集》（*Collected Papers of Srinivasa Ramanujan*，纽约：切尔西出版公司，1962 年），第 xi 页。

[12] 这种推导欧拉公式的精彩方法见于 L.W.H. Hull 的笔记“Convergence on the Argand Diagram”，《数学公报 43》（*Mathematical Gazette* 43，1959 年），第 205~207 页。George W. Hart 指出了此点，并建议采用不同的字体，在此致谢。

[13] Herbert Turnbull 引用 Felix Klein 的《数学的世界》（*The World of Mathematics*，纽约：Simon and Schuster，1956 年）中的“The Great Mathematicians”，第 1 卷，James R. Neuman 编，第 151 页。

[14] 不过这并不是通用的表达式。例如，数学家有时会争论 π 的定义是不是最经济的。也就是说，对数学和科学中业已发现的 2πs，以及把 π 作为单位圆的弧度所产生的极大简化，此处把圆的周长与半径的比值定义为基本常数的做法是否更美、更经济呢？换句话说，有没有什么例子能表明 π 确实有其美和经济性呢？$e^{i\pi}+1=0$ 就是最明显的例子。这个方程如果变成 $e^{i\pi/2}+1=0$ 的话，乍看上去美感会降低。不过数学家发现了一个窍门。假定用 ψ 表示 2π，那么就能写出一个更美、更经济的公式：$e^{i\psi/n}=\sqrt[n]{1}$ 。欧拉公式只是上述公式的特殊情形。因为 1 有一个平方根是−1，所以公式 $e^{i\psi/n}=\sqrt[n]{1}$ 更为一般化。欧拉公式是 $e^{i\psi/n}=\sqrt[n]{1}$ 的特殊情形，就像毕达哥拉斯定理是余弦定律的特殊情形一样。

[15] Larry Wilmore，引自《纽约时报》（*The New York Times*），2007 年 4 月 15 日，第 4 节，第 4 页。

[16] Len Fisher，“日常生活中的方程”，《新科学家》（*New Scientist*），2005 年 7 月 30 日；Simon Singh，“谎言、该死的谎言和 PR”，《新科学家》，2005 年 8 月 20 日。

[17] 关于此点的讨论，参见 William Steinhoff，《乔治·奥韦尔与 1984 的起源》（*George Orwell and the Origins of 1984*，安阿伯：密歇根大学出版社，1975 年），第 12 章。

[18] 这位领导人指的是 Eugene Lyons，他起草过前苏联的第一个五年计划，引自 Steinhoff，《乔治·奥韦尔与 1984 的起源》，第 172 页。

[19] 引自 Robert A. Orsi，“2+2=5”，《美国学者》（*American Scholar*，2007 年春）76，第 34~43 页。

5

科学上的莎士比亚剧
热力学第二定律

$$S' - S \geqslant 0$$

说明：整个世界的熵值向着达到最大值的方向变化。
发现者：各国科学家。
发现时间：19世纪40年代到19世纪50年代。

我经常会遇到很多这样的人：按照传统文化的标准来看，他们算得上是受过高等教育的人，这些人喜欢兴致勃勃地表达自己对科学家的怀疑。有那么一两次，我被激怒了，就问他们谁能描述一下热力学第二定律。结果一下子冷场了，没有人能说得上来。在科学界来看，我问的这个问题其实就相当于问人文领域的人："你读过莎士比亚的作品吗？"

——C.P. 斯诺（C. P. Snow），《两种文化》（*The Two Cultures*）

热力学第一定律和热力学第二定律的叙述很简单。热力学第二定律是由鲁道夫·克劳修斯提出的。他发明了"熵"（entropy）一词来描述混乱度。克劳修斯是这样描述这两个定律的："世界的能量总量是恒定的，世界的熵值总是向着达到最大值的方向变化。"简而言之就是："你永远不可能赢，也不可能不赔不赚。"本文一开始给出的公式是马克斯·普朗克提出的，其中 S 是之前的熵，S'是一段时间之后的熵。世

界上的所有活动都会涉及热力学第二定律。不理解这一点，就无法理解世界运行的内在机制。C. P. 斯诺认为，问人家能否描述热力学第二定律就好比是问："你读过莎士比亚的作品没有？"当然，这也正是他本人提问的初衷。对于自认为有文化的人来说，这两个问题只要有一个答不上来，就可算是颜面尽失了。

谈到这个话题，我的想法更偏激一些。我认为热力学第二定律实际上就像一部莎士比亚的作品。它也有自己的故事，故事里面还有精心刻画的角色。它对人的生活有着重要的意义。故事情节的展开，在某种程度上与莎士比亚剧如出一辙。

以下就是"剧情介绍"。

序幕

18 世纪末的欧洲

此时一种新的力学理论已初见端倪。蒸汽机和其他技术的发展，使人们把注意力集中到与热有关的现象上来。受实际需要和好奇心的驱动，很多发明者开始尝试设计更好的蒸汽机。不过当时人们对热几乎一无所知，这些人的工作也只不过是"缝缝补补"。人们似乎认为热就是力，热可以用来做功。但是，热却不能用牛顿的推力和拉力来解释。此时，人们对热的理论产生了迫切的需求，不久就出现了一种称为"热质说"的粗糙理论。该理论于 18 世纪下半叶由法国科学家、"现代化学之父"安托万·拉瓦锡（Antoine Lavoisier）提出。热质论认为热是不断流动的、肉眼看不见的流体，没有重量。这是人们把热理解为力的雏形。随后有几位科学家，受好奇心和科学家的自尊和自信的驱使，也开始将注意力转移到热力上来。很快，他们就在热力的利用过程中是否涉及热的守恒和转化这一问题上，发生了争执。也就是说，热的总量是否守恒？或者热能否转化成其他能量？这场争论的解决成为新力学的关键。

剧中人物

路德维希·玻耳兹曼
（Ludwig Boltzmann，1844—1906 年）



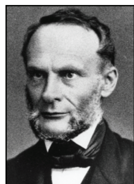
玻耳兹曼是奥地利物理学家，情绪容易波动，患有抑郁症。他采用统计方法说明了许多微小的原子如何能产生物质的宏观属性。他因提出玻耳兹曼方程 $S=k\log W$ 而成名。玻耳兹曼的原子理论和自己的工作遭到了声名显赫的同行攻击，他为此深感苦恼。1906 年，与家人在意大利的的里雅斯特附近度假的玻耳兹曼，趁妻子和女儿游泳时，上吊自杀。他的墓地在维也纳，碑上刻着玻耳兹曼方程。

赫尔曼·冯·亥姆霍兹
（Hermann von Helmholtz，1821—1894 年）



亥姆霍兹是德国物理学家，精通多个领域并多有建树，包括声学、美学、解剖学、生物学、磁学、数学、力学、气象学、眼科学、光学、现象学、哲学、物理学、生理学和心理学。他发明了检查肉眼用的检眼镜。许多科学名人都是他的学生，如诺贝尔奖获得者阿尔伯特·迈克尔逊（Albert Michelson）、马克斯·普朗克（Max Planck）和威廉·维恩（Wilhelm Wien）。

鲁道夫·克劳修斯
（Rudolf Clausius，1822—1888 年）



克劳修斯是德国物理学家。他指出热力学两大定律是统一的，从而平息了转化论者和守恒论者之间的纷争。一个涉及的是热和机械运动在转换中某物的守恒（不久后称能量守恒），另一个涉及的是热向能的转化。克劳修斯首次采用了“熵”（entropy）一词，用 S 表示。他也声称自己是热力学第二定律的发现者。

萨迪·卡诺
（Sadi Carnot，1796—1832 年）



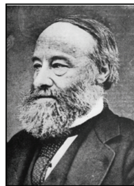
萨迪·卡诺是一位工程师，喜欢安静。父亲拉扎尔死后，他搬进了父亲留给他的公寓里。萨迪对提高热机效率很有兴趣，写了一篇颇具影响力的论文——《关于热动力的设想》（*Reflections on the Motive Power of Heat*）。论文由他的兄弟希波吕忒编辑。论文中包括了守恒和可逆的思想，以及著名的“卡诺循环”。不过它却没有引起人们的重视。从此以后，卡诺没有再发表任何论文。他因罹患猩红热、脑膜炎和霍乱，死在精神病院，死时年仅 36 岁。

拉扎尔·卡诺
（Lazare Carnot，1753—1823 年）



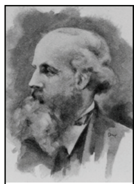
拉扎尔·卡诺是法国军事工程师，擅长废物处理。因为服兵役的关系，卡诺中断了对水力机效率低下问题的研究。后来，他因为勾引一个已订婚的女人而被捕入狱。法国大革命时，卡诺屡立奇功，绰号“胜利组织者”（Organizer of Victory）。他育有两子：萨迪和希波吕忒。两个儿子由他亲自教育。

詹姆斯·普雷斯科特·焦耳
（James Prescott Joule，1818—1889 年）



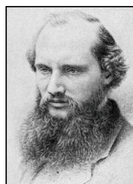
焦耳年轻时就在父母的啤酒厂搭建了一个家用科学实验室。几年后，他在实验室中对热能与电能、机械能和化学能之间的相互转化进行了高精度的测定。他首次测出了热功当量。他的工作推进了能量转化的思想，也解决了转化论者和守恒论者之间的纷争。

詹姆斯·克拉克·麦克斯韦
（James Clerk Maxwell，1831—1879 年）



麦克斯韦是个地道的神童，但在学校里却经常因为衣着寒酸、乡下口音重和问的问题直白，而遭到同学们的嘲笑。他们还给他起了个绰号叫“蠢货”。麦克斯韦空前出色地使用类比的方法创建了电磁学领域，奠定了电子时代的基础。他的其他成就还有：解释了土星环、气体的行为、陀螺的原理，制造出了“有史以来最棒的陀螺”。麦克斯韦 48 岁离开了人世。

威廉·汤姆逊
（William Thomson，1824—1907 年）



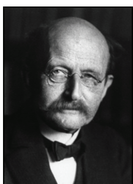
汤姆逊是一位数学教授的儿子。他知识渊博，通晓三门语言，是一位具有远见卓识的年轻人，后来成为开尔文勋爵。他对转化论者和守恒论者之间的争论感到苦恼，决心解决这一问题。汤姆逊提出了新热动力学，并称之为热力学（thermodynamics）。他是《自然哲学教程》（*Treatise on Natural Philosophy*）的作者之一，该书堪称是热力学方面的牛顿《原理》。他也称自己是热力学第二定律的发现者之一。

拉姆福德伯爵
（Count Rumford，1753—1814 年）



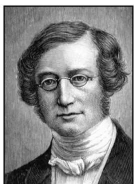
拉姆福德伯爵是个英国兵痞、业余科学家和间谍，周旋于富有的寡妇之间。他进行了一些热学实验，对热质说进行了反驳。热质说是拉姆福德最后一个情妇的前夫拉瓦锡提出的。拉姆福德认为热不是物质，而是来自于摩擦产生的运动，并采用这一思想比较了各种不同类型的功。他认为自己是牛顿再世。

马克斯·普朗克
（Max Planck，1858—1947 年）



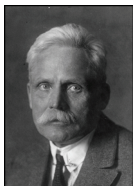
普朗克不顾老师的警告（即物理学不会有新的发现了），全身心地对已有理论（热力学）进行整理。他发明了量子论，改变了整个世界！他的大儿子在一战中死于凡尔登，二儿子因为在二战期间参与谋杀希特勒而上吊自杀。普朗克于 1919 年获得诺贝尔奖。为纪念他对科学作出的卓越贡献，一家世界闻名的研究机构改以他的名字命名，同样还以他的名字命名了一颗周长 43 公里的小行星。

罗伯特·迈尔
（Robert Mayer，1814—1878 年）



迈尔在东印度的一艘荷兰船只上做医生时，发现船员的血异常地发红。这说明血液中的氧很多。他认为这是人在热带地区新陈代谢变慢的缘故，并推断机械功和热是可以互换的。在这方面他写出了一篇晦涩难懂的论文，被杂志拒稿。之后他又对该论文进行了修改，在其他地方发表。迈尔称自己发现了热力学第二定律，但这一说法未被大家承认。抑郁至极的迈尔从三层楼的窗户跳了下去。最后迈尔被送进了收容所。

威廉·维恩
（Wilhelm Wien，1864—1928 年）



维恩自认为是一个农民，物理只是他的第二职业。他提出了维恩定律，该定律利用热力学第二定律，得出了高温条件下温度对辐射的影响，从而将人类领进了“量子力学的大门”。他发现了一种带正电子的粒子。后来的研究发现，这种粒子就是今天所谓的质子。维恩 1911 年获得诺贝尔奖。火星上有一个直径 120 公里的坑就是以维恩的名字命名的。

第 1 幕

巴黎和慕尼黑，18 世纪末期

第 1 场：巴黎，1803 年

拉扎尔·卡诺是一位军人出身的工程师，在发现并解决行政效率和机械效率低下问题方面天资甚高。他发表了一篇关于水力机的论文《平衡和运动的一般原理》（*General Principles of Equilibrium and Motion*）。卡诺从水出发，进一步写道："水所能产生的最大能量与落差有关。"除调查和清除废物源之外，他还考虑如何提高水力机的效率。因服兵役，他没能将这项工作继续下去。后来，他因为勾引一个已订婚的女人而被捕入狱。法国大革命爆发后，卡诺出狱，参加了革命。因在部队调遣、训练和补给方面屡立奇功，卡诺被人们称为"胜利组织者"（Organizer of Victory）。他有两个儿子，都是在家由他亲自教育长大。他们果然也继承了父亲的传统：第一个儿子叫萨迪（用了一个波斯诗人的名字），是一名军人出身的工程师；另一个儿子希波吕忒，是一名记者和政治家。

第 2 场：慕尼黑，1797—1798 年

拉姆福德伯爵是慕尼黑的一个兵痞，也是一位业余科学家。他经常周旋于有钱的寡妇之间。另外，他热切地想要了解热的奥秘。他把 6 磅（1 磅=0.454 千克）重的黄铜炮身放到一桶水中，再插进一个钻头，钻头连着绞盘。之后，拉姆福德让马转动绞盘，他发现 2.5 小时所产生的热量足以烧开桶里的水。拉姆福德于是宣布拉瓦锡（拉姆福德一个情妇的前夫）的热质说是错误的。因为通过这种过程所产生的热似乎是无穷无尽的。热既不是从铜中产生，也不是从水中产生的。很显然，热是由钻头和炮筒之间摩擦的某种运动形式产生的。他数了一下将同样量的水烧开需要多少支蜡烛，对热和机械力的大小进行比较，并将这一结果报告给了皇家学会，说热学定律的重要性不亚于万有引力定律，暗示自己是

牛顿再世。但拉姆福德终究不是牛顿。他的论述并不能完全令人信服，并且他也没有形成完整的热学理论，只有一些启发性的实验结果。不过他的思想，例如对可产生同样热量的各种类型的功（蜡烛、马）进行定量的比较，以及热通过不同形式转化成的功，引发了能量转化和守恒之间的对立。

第 2 幕

巴黎、曼彻斯特和牛津，19 世纪 20 到 40 年代

第 1 场：巴黎，1823 年

萨迪·卡诺是一位工程师，喜欢安静。父亲拉扎尔死后，他搬进了父亲留给他的公寓。萨迪决心要将父亲的工作继续下去，并开始撰写一篇题为《关于热动力的设想》(*Reflections on the Motive Power of Heat*)的论文，考虑如何提高蒸汽机的效率和实用性。卡诺怕自己的文字深奥难懂，难以激起普通读者的兴趣，于是请自己的兄弟希波吕忒对稿子进行编辑，使文字顺畅易读。他是这样写的：蒸汽机“似乎注定要在文明世界中掀起一场革命”。但是，他又继续写道：“可人们对蒸汽机的理论还知之甚少。”要提出蒸汽机理论，必须要从如何最有效地利用蒸汽机这个一般问题入手。卡诺认为，研究蒸汽机，最重要的是要研究蒸汽机的最大工况，即最大输出。例如，蒸汽机中一定质量的煤能将一定量的水提升多少度？接下来他就写热。热机中的热量，就像是抽水机中的水，热量在从高温物体传递到低温物体的过程中是守恒的。热机的最大效率取决于温差的大小。效率最高的热机可通过一种理想循环来实现。该循环包括膨胀和压缩过程，热机在其中可逆地工作。热量在两个高低温度之间的往复过程中保持守恒，不会因为摩擦和耗散而转移（浪费)。对于热机而言，这是最重要的一点。但是《设想》一文却几乎把它完全忽略了。此后，卡诺没有再发表论文。他因罹患猩红热、脑膜炎和霍乱，死在精神病院，死时年仅 36 岁。

第 2 场：曼彻斯特，19 世纪 40 年代

焦耳小时候就在父母的啤酒厂搭建了一个实验室。他想在实验室中对热能与电能、机械能和化学能之间的相互转化进行高度精确的测定，例如，螺旋叶片搅动水所产生的摩擦会导致水温的升高。他测出了热的机械当量：772 英尺磅[①]的功可以使 29 立方英寸的水的温度升高 1 华氏度。

第 3 场：牛津，1847 年

能量的守恒（卡诺的方式）和转化（焦耳的方式）之间的冲突已经白热化。年轻的威廉 · 汤姆逊是一位数学教授的儿子。他知识渊博，通晓三门语言，是一位具有远见卓识的年轻人。他在巴黎发现了唯一一篇已发表的关于萨迪 · 卡诺著作的评论。他被这篇评论深深吸引，于是试着去找卡诺的原著，没有找到。之后他参加了在牛津举办的一个会议，听了焦耳的报告。会议组织者对待焦耳很不友好，要他长话短说，不要啰嗦。不过，焦耳的话却使汤姆逊颇为震撼。如果卡诺那壮观的研究成果所依据的事实就是热机中热的总量是不变的，那热量怎么可能转化成其他能量呢？汤姆逊认为，焦耳的研究当中一定有“很大的纰漏”，他要把这纰漏找出来。

第 3 幕

英国和德国，19 世纪 40 到 60 年代

第 1 场：格拉斯哥

这次，“坚信卡诺的守恒理论是正确的以及焦耳的研究一定存在问题”的汤姆逊又受到一次震撼。他读了德国物理学家鲁道夫 · 克劳修斯的一篇论文。克劳修斯发现了卡诺和焦耳方法上的冲突，他仔细研究了热运动理论：根据该理论，热和气体都是由不断运动的微小粒子组成的。

---

① 英尺磅为功的单位，1 英尺磅=1.355 牛顿米。——译者注

克劳修斯认为卡诺和焦耳之间的分歧只是表面上的，实际上并无矛盾，两者是一致的。一个涉及的是热和机械动作在相互转化过程中某物的守恒（不是热，之后不久被命名为能量）；另一个涉及的是热能向能量的转化，以及热能不能自发从低温物体传递到高温物体的属性。汤姆逊受此启发，开始跨越性地将克劳修斯的研究与新热力学（heat-mechanics）联系起来。1854 年，汤姆逊将这门学科命名为热力学（*thermodynamics*），取希腊语中“热”和“力”之意。汤姆逊写道：“所有热机中都有一部分热量，会不可逆转地损失掉，也就是‘浪费’掉了，虽然热量本身并没有消失。”这就是汤姆逊版本的克劳修斯的热力学第二定律。克劳修斯于 1865 年写出了一系列的论文，研究工作达到了顶峰，他把能量自发转移的趋势（现在的说法是混乱度）命名为“熵”，取希腊语“转化”之意。他用 S 表示体系的状态函数——熵，并使用公式 $\int \mathrm{d}Q/T \leqslant 0$ 。1867 年，汤姆逊与合作者泰特（Tait）写出了《自然哲学教程》（*Treatise on Natural Philosophy*）一书，该书堪称热力学领域的牛顿《原理》。1872 年，克劳修斯提出了后来所谓的热力学第一和第二定律——“世界的能量是守恒的”，“整个世界的熵值向着达到最大值的方向变化”。

第 2 场：海尔布伦，德国

发现的先后之争爆发了。1847 年，德国物理学家罗伯特·迈尔在读到焦耳关于热能和机械能相互转化的论文后，称自己是第一个发现该规律的人。7 年前，迈尔在东印度的一艘荷兰船只上做医生时，发现船员的血异常地发红。这一现象表明血液中富含氧。迈尔认为这是由于人在热带地区代谢变慢的缘故。受此启发，他写了一篇机械能与热能相互转化的论文，寄到了德国顶级科学期刊《物理学和化学年鉴》（*Annalen der Physik und Chemie*）。可惜论文写得太差，编辑认为太过不切实际，没有给他答复。无奈的迈尔只能修改后在别处发表。后来迈尔与焦耳就发现的先后问题展开了激烈的争论。抑郁至极的他从三层楼的窗户跳了下去，

最后被送进了收容所。与此同时，另一位德国物理学家赫尔曼·冯·亥姆霍兹因为在1847年发表了一篇关于“力的守恒”的论文，也成为发现热力学第一定律的有力竞争者。就谁最先发现热力学定律这一问题，泰特和克劳修斯相持不下，两人在各种期刊和书籍中互相攻击。

第4幕

伦敦、格拉茨和维也纳，19世纪70年代

第1场：伦敦和格拉茨

此时，另一场纷争也爆发了。这次纷争的主题是热力学第一、第二定律哪一个更重要。这两个定律看上去是相互矛盾的。第一定律（热/能量守恒）表明过程是可逆的，也就是说无法区分物理过程的“始态”和“终态”，两者可以相互转化。而热力学第二定律说的却是不可逆性（热不能再次完全转变为功），即后来所谓的“时间之矢”。也就是说变化是向着某个方向发展的。这个问题在克劳修斯的特长——气体动力学理论中渐趋成熟。气体是一个“大物体”，它受到不可逆过程和热力学第二定律的控制，但它又是由“小物体”——原子和分子组成的。原子和分子遵循热力学第一定律控制的可逆牛顿原理。1859年，麦克斯韦无意看到了克劳修斯关于气体动力学理论的论文，于是就认为他把土星环当作许多小天体来研究的方法可能也适用于气体的研究。麦克斯韦认为气体中相互碰撞的分子并不能达到速度都相同的平衡状态，相反，气体的速度应该分散在某一特定值附近。想象一下车站里走来走去的人群：人们走动的速度并不是严格相同的，但是大部分人的速度基本上是相同的，只有少数人走不动或者走得很快。进而，理解气体的行为，亦即理解大量气体分子的行为时，对各个分子的位置和速度的追踪是没有必要的，只要知道速度和动量的分布就够了。麦克斯韦只用统计方法和牛顿力学假设，就提出了一个描述气体分子速度范围的方程。这个方程的曲线呈钟

形：中间平均速度附近的分子多，两头的分子少（几乎没有速度或者速度很快），偏离平均速度越远，分子数就越少。不过，克劳修斯在 1865 年发表的论文和麦克斯韦本人的实验工作，都迫使他对理论进行修改。1867 年，麦克斯韦发表了修改后的理论。他得出结论：热力学第二定律只是统计意义上的，只有在粒子数很多的情况下才适用，它不适用于单个粒子的运动。他写道："如果把一杯水倒进大海里，那么就再也不可能把这杯水从海里取出来了（也就是无法使杯子中的水分子与最初的完全一样）。"[1]热力学第二定律的描述与此类似。麦克斯韦认为：在原子尺度上，可逆是可以实现的，所以热力学第二定律并不成立。但是为什么可逆原则上在大型物体中无法实现呢？为什么热量不能从低温物体传递到高温物体？1867 年，他写了一封信给泰特，信中有一个很有意思的假想实验：假定有一个小妖，可以找出箱中气体里运动速度较快的分子，这样一来，在适当的时间控制一个阀门的开关，就可将速度较快的分子集中到箱子的一侧，从而使热量传到箱子的这一侧。似乎这个假想的生灵就可以通过这种方式使热量从低温物体传递到高温物体，从而反驳了汤姆逊的耗散理论。1871 年，麦克斯韦在《热学原理》（*Theory of Heat*）一书中以"热力学第二定律的局限"（Limitations of the Second Law of Thermodynamics）为题，用一小段文字阐明了上述思想。事情到此为止好像是结束了，热力学第二定律只是个统计问题。

第 2 场：格拉茨和维也纳，19 世纪 70 年代

路德维希·玻耳兹曼进一步扩充了麦克斯韦的工作。1868 年，也就是麦克斯韦论文发表后的第二年，玻耳兹曼提出了一个对任何类型气体都适用的气体分子能量分布表达式。为推出该表达式，他作出一个重要的假设，即所谓的能量均分定理。按照该定理，分子中储存的能量是均匀地分布在所有自由度上的。该工作还涉及一个著名的量：玻耳兹曼常数。我们现在用 k 来表示该常数，k=1.38 × 10^{-23} J/K。这个结果完全是从

统计学角度对热力学进行解释得到的。1872 年，玻耳兹曼进一步对该工作进行了研究，写出了一篇具有革命性影响的论文。不过论文的题目倒很普通——“气体分子热平衡的进一步研究”（Further Researches on the Thermal Equilibrium of Gas Molecules）。他在论文中推导出了一个与熵有关的函数，现在称为 H 函数。该函数表明，熵的值总是随时间不断增加，直到达到最大值为止。玻耳兹曼以全新的方式证明了热力学第二定律，明确说明了不可逆性，以及熵是如何随时间变化的。不过这一工作却遭到了友善的攻击：一方面是汤姆逊，他在 1874 年发表的论文中提到了麦克斯韦的小“生物”，并称其为“小妖”；另一方面是玻耳兹曼的前导师约瑟夫 • 劳施密特（Josef Loschmidt）。劳施密特在 1876 年指出热力学第二定律与热力学第一定律的关系中涉及的一些谜团仍未解开。哪怕是最复杂的多体系中，星星在太阳周围的位置也是循环往复的，不断重复着相同的图式。为什么在热力学体系中这点就不能成立？还有，如果将两种气体混合，按照 H 曲线，熵是增加的；但是如果改变所有气体分子的速度方向，那么 H 曲线（“时间之矢”）是不是就要颠倒过来，从而违反第二定律呢？玻耳兹曼（1877 年）答复说，如果某个大状态对应于许多可能性相同的小状态，那么大状态的概率取决于小状态的数量。玻耳兹曼是明确采用概率的方式解释熵的。他将概率引入电磁学中，证明了不可逆性在热力学中的中心地位。

牛顿定律+众多物体构成的大物体+概率论=时间之矢

完全不可能的事只能说不太可能。在大尺度下，比如玩骰子，发挥作用的就是统计规律了。1879 年，玻耳兹曼以前的老师斯蒂芬对该工作进行了进一步的扩充，提出了斯蒂芬−玻尔兹曼定律。该定律指出黑体辐射取决于黑体的温度。但是，晚年的玻耳兹曼，因为个人和事业上的一些挫折，变得极为抑郁。1906 年，他在意大利的的里雅斯特附近度假，趁妻子和女儿在外面游泳时，上吊自杀。他的墓志铭上刻着的就是玻耳兹曼方程。不过方程的形式不是最初他写的那个样子，而是经过马克

斯·普朗克改写的 $S=k\log W$。

第 5 幕

柏林，19 世纪 90 年代

第 1 场：柏林，19 世纪 90 年代早期

物理学家威廉·维恩是一个内向的人，他想跟父母一样当个农民，但却屡受阻挠。他扩充了玻耳兹曼的热力学第二定律的思想。维恩在皇家标准局的联邦物理技术研究院（Physikalisch-Technische Bundesanstalt）工作。他与其他几位科学家，出于理论上的兴趣和实际需要（电灯校准），共同研究“黑体辐射”。某个物体如果能吸收所有投射到其上的辐射，就称为“黑体”。如果对黑体进行加热，它就会发热发光。根据经典力学，物体中存在一个类似于谐振子的东西。它就像一个小天线，以电磁波的形式吸收和释放能量。电荷好像是位于不同弹性的弹簧之上。弹簧来回振动的频率取决于弹簧自身的弹性，振动强度由温度决定。麦克斯韦彻底解释了辐射的产生、吸收和传播。通过测定辐射，实验学家绘制出了各个温度下波长和强度的曲线。基于斯蒂芬和玻尔兹曼的工作，维恩写出了一篇题为“黑体辐射与热力学第二定律的新关系”（A New Relationship Between the Radiation from a Black Body and the Second Law of Thermodynamics）的论文。论文中提出了维恩定律。该定律利用热力学第二定律得出高温条件下，温度对辐射的影响：“在黑体的正常发射波谱范围内，波长的移动和温度的变化是有一定规律的：二者的乘积总是常数。”[2]这一规律被称为“位移定律”，它于 1896 年被再次提出。同样是在 1896 年，皇家标准局的科学家们制作了一个特殊的烘箱，用它测定辐射的波长。测量的范围集中于容易检测的短波长区域。虽然在能量降到足够低时，用经典理论就能完全解释清楚，但实验学家发现，随着能量曲线向越来越长的波长区域延伸，维恩定律便不再成立。因为维恩定律

直接依赖于经典物理学的逻辑架构，所以这一现象值得引起注意。

第 2 场：柏林，19 世纪 90 年代晚期

马克斯 • 普朗克是一位“不情愿的革命者”，最得意的课题是热力学。此时他也将注意力转移到黑体辐射问题的研究上来。1878 年的普朗克还是一名研究生，他偶然发现了克劳修斯的一些论文，被深深吸引。此后普朗克就开始写论文，对热力学第二定律的已有叙述进行评论。那时，人们都认为热力学几乎已经是完备的了，没有什么激动人心，或者被人们看好的领域可供年轻的科学家挖掘。但普朗克的性格比较保守，他对理论基础的巩固很感兴趣，而且玻耳兹曼对于热力学第二定律的解释还有让他不解的地方。他认为，定律应该是绝对的，没有任何例外；热力学第二定律应与热力学第一定律一样具有普遍意义，而不是靠统计学的戏法变出来。1895 年，普朗克来到柏林。他的助手策梅罗提出，热力学第二定律永远无法证明，不仅如此，不管多么复杂的机械系统，最后都可以回到初始状态。“理论物理目前所面临的最重要的（问题）”就是热力学第二定律的概率性和不可逆性与牛顿力学的不变性和可逆性之间的矛盾。而黑体辐射似乎是解决问题的关键，答案可能就在谐振子吸收和释放能量的方式上。当时的柏林正是黑体辐射的研究中心，维恩和其他几个实验学家都在柏林。普朗克或许正好可以利用他们的工作，表明如何将电磁理论与热力学定律联合起来，从而解释平衡状态下的辐射分布。普朗克从修正玻耳兹曼的工作以使其清晰明了和用频率代替波长重写维恩定律开始，试图将热力学、统计力学和电磁理论统一起来。1897 年，普朗克在普鲁士学会作了系列报告（时间跨度 7 年）的第一讲，题目是“论不可逆辐射过程”（On the Irreversible Radiation Processes），旨在解决他所谓的“理论物理的根本任务”——热力学两大定律的统一。因为这两大定律存在矛盾，他一开始就指出研究黑体辐射的迫切需要。热力学第一定律，或“能量守恒定律”认为，诸如摩擦之类的任何效应都可以

在微观上分解成机械的可逆过程。但是热力学第二定律，也就是“熵增原理”却要求“自然界中的所有变化只能向着一个方向进行”。他接着告诉听众：“两大热力学定律的统一是理论物理的根本任务。”几次报告之后，在 1900 年 10 月 19 日，普朗克提出了一个经验公式。这个公式跨越了维恩定律适用的高能区，经典力学适用的低能区，以及位于高能区和低能区之间的、维恩定律与实验数据符合得不是很好的中间区域。普朗克说，这一公式是从“谐振子不能在固有频率下振动，而只能在特定频率（参数 h）下振动”这一思想出发的。所提出的熵表达式完全是主观上的。一向非常谨慎的普朗克说，“就目前来看”，该工作“与已发表的数据吻合得最好的光谱方程一样令人满意”。他得出结论说：“因此，我应该将你们的注意力转移到这个新公式上来。在所有的公式当中，它的形式是最简单的。”[3]接着他就把量子力学的思想引入到了物理学中。不过他这么做时还是很犹豫的，甚至是不情愿的。当天晚上，就有一位实验学家受到鼓舞，回到实验室去验证普朗克的“新公式”，实验结果与公式相符。普朗克非常激动，继续投入到工作中。“在度过了这辈子最辛苦的几个星期之后，黑暗褪去了，一种之前从未想象过的前景开始展现出来。”这就是说，普朗克的的确确将热力学、电动力学和经典力学结合起来了，并对实验中最后一个令人困惑的问题进行了解释。这种结合简洁漂亮，意义也不容置疑。普朗克和当时参与的其他所有人都没有想到，热力学基础的建立会导致另一种新能量概念的出现，引领人们进入一个全新的世界。

尾声

乌云

1900 年 4 月，汤姆逊在皇家学院发表了一次演讲，题为《在热和光动力理论上空的 19 世纪乌云》（*Nineteenth Century Clouds over the Dynamic Theory of Heat and Light*）[4]。汤姆逊说：热和光理论（热力学和电磁学）的“美丽和清晰”是 19 世纪科学的最高成就，但这个胜利却

因“两朵乌云的遮盖而变得模糊起来”。

第一朵乌云就是很难理解地球如何在以太中运动。19 世纪初期的科学家认为以太存在于固体中的原子之间。它就像“吹过树林的风”一样“吹过”原子。但麦克斯韦认为以太一定更类似于液体或者固体，而且以太能对通过其中的物体施加力的作用。这意味着地球相对以太的运动是可以探测到的。不过汤姆逊接着又说，阿尔伯特·迈克尔逊（Albert Michelson）和爱德华·莫雷（Edward Morley）最近进行了一个“绝妙的实验”，排除了上述可能。该实验在设计和实施上都无懈可击。此时，乔治·菲兹杰拉德（George Fitzgerald）和亨得里克·洛仑兹（Hendrik Lorentz）独立提出了一个解决方案：如果将在以太中运动的物体运动方向上的维度稍作变化，科学家们就可以保住“以太”，而且只要一亿分之一的变化量（地球绕太阳的运行速度与光速之比的平方）就够了。汤姆逊虽然认为这种想法“非常聪明”，但同时也难免有怪诞之嫌。他得出结论：“恐怕我们依旧得认为第一朵乌云的密度是很大的。”

第二朵乌云与“能量等分定理”有关。该定理由麦克斯韦和玻耳兹曼提出。根据该定理，分子中的能量均匀分配在各条可能的路径上。它解释了人们熟知的固体在高温下的热容量定律问题，但却与固体、气体和金属在低温下的实验结果存在严重的矛盾。此时热力学已经取得了极大的成功，所以当时的科学家对这种矛盾非常困惑，想要找到一个解释。汤姆逊承认自己对此无能无力。他引用了英国物理学家瑞利爵士的一句大胆的话：我在等待能把我们从能量均分定理的“破坏性的简单性中解放出来”的新原理出现。汤姆逊在报告的结论中称：第二朵乌云“在 19 世纪的最后 25 年中，遮蔽了光热分子理论的光辉”，如果有朝一日果真有新原理出现，第二朵乌云就会散去。

汤姆逊不会知道，正是 19 世纪的这两朵乌云，不久之后演变成了 20 世纪的飓风，也就是“相对论”和“量子力学”。

其他版本的热力学第二定律戏剧在细节和广度、角色的数量和大小

上与本书的这一版本可能会有不同。但我要说的是，这部剧是莎士比亚式的。剧中的演员都是伟大的人物，他们倾其身心，对科学研究满怀执着。他们对实验结果和实验预期的偏差感到苦恼和困惑，要进一步进行研究，来更好地了解世界。这种苦恼和困惑有时是深刻的、悲剧式的，而剧中的每一幕都以此开场。试想一下，世上的剧中，还有哪一部能有如此独一无二、精巧刻画的人物？还有哪一部对人类和世界产生的影响能比它更深远？



茶　歇

不可能的科学

> 科学上的每一次进步都是要付出代价的。人类在智力上所取得的每一个新的成就都意味着对旧有观点和概念的抛弃。因此，这么说来，科学家们所自恃的对自然的“理解”将随着知识的增加而不断减少。
>
> ——海森堡

科学上的许多定理都是如下的形式：“如果这么做，结果就会如何如何。”牛顿第二定律就是一个例子。它说的是某个物体的加速度与所受力的大小成正比。诸如此类的原理表明某些效应实际上是有可能发生的。然而，有一小部分却是属于另一类，这类原理说的是“哪些是不可能发生的”。这些原理表明某些效应在物理上是不可能的。

关于后者，众所周知的例子就是热力学第一和第二定律，其他的例子还有海森堡的不确定性原理和相对论。相对论的基本思想是认为绝对速度是不可能的，以及任何物体的速度都不可能超过光速。这些原理所代表的并不是“新物理学”，而是根据其他原理得出的推断性结论。这些

原理的不同之处在于它们的形式。而这种形式（断言某物在物理上是不可能的）促使科学家们去突破。

不可能科学有好几个名字。一个是“忘了它”科学，另一个是“没门”科学。半个世纪以前，数学家和科学史学家埃德蒙·惠特克（Edmund Whittaker）爵士提出了“无效假设”，断言“有些事情是无法实现的，哪怕可能实现的道路有千万条也无济于事”。

惠特克写道：“无效假设并不是某个实验，也不是有限个实验的直接结果。它不提测量、数值关系或者解析方程，只是断言一个结论，即某件事，不管怎样努力地去做，都是注定要失败的。”

无效假设既不像人们从经验中发现的实验事实，也不像先验的、按定义正确无疑的数学描述。惠特克接着说：但是这些假设是科学的基础。他说：热力学就可以认为是从无效假设中得出的一组推论，即能量和熵的守恒。他接着说：在遥远的将来，科学上的各个分支完全可能像欧几里得的《几何原本》那样，以合理的不可能做到的假设呈现出来。

然而，“没门”科学之所以重要还有另一层的原因：它强烈吸引着持相反意见者。我所说的“持相反意见者”不是那些白费心思，想要造出永动机，来推翻热力学第二定律的那些傻子愣子。我指的是发现“没门科学”的漏洞，对它提出挑战的具有严肃认真态度的科学家。正是在寻找这些漏洞的过程中，他们阐明了科学领域的基础。

持相反意见的科学家不仅在科学发现上，而且在不确定性原理等科学原理的解释上也发挥了作用。1926年，海森堡正推广他的新矩阵力学（完全以正规方式研究原子物理），宣称物理学家必须要摒弃在原子内部电子上观察到宏观属性（位置、动量、时间和空间等）的幻想。帕斯卡·约尔当（Pascal Jordan）就担当了持有相反意见的角色。他提出了一个假想实验，来反驳海森堡的论点。约尔当说，假定可以将显微镜的温度降低到绝对零度，那么就可以测量到原子和原子内的组成部分的准确位置和动量。海森堡似乎受此启发，开始考虑观察设备与观察到的现象之间的

相互作用。以此为契机，不久后海森堡就提出了不确定性原理。在这里，约尔当，这位意见不同者，强迫海森堡从操作上而不是从哲学上来思考问题，并最终使海森堡澄清了观察结果中涉及的物理学原理。

此后，爱因斯坦也担当过意见不同者的角色，颇为引人注目。他反对的对象是玻尔。爱因斯坦的目的是想出可以准确测定出粒子的位置和动量的巧妙办法。他的尝试虽然失败了，可是由此所引发的讨论却极大地促进了物理学家对量子力学的本质和意义的理解。

物理学上另一个著名的意见不同者是麦克斯韦。他提出了一个假想实验，在密闭的盒子内，设置一个小妖，来控制盒子内的一扇小门。通过门的开关，这个小妖就可以将盒子一侧速度较快的分子移动到另一侧，从而使热量定向流动，违反热力学第二定律。对这个假想实验进行的讨论帮助人们澄清了热力学当时的概念之谜。

海森堡说，科学上的每一次进步，都意味着科学家们降低了对自然的“理解”主张。这番话自然有夸大之嫌。当然，科学的进步更多地是提出更加细致、复杂的概念，来替换、包含已有的简单概念。但是，这些更加细致和复杂的概念常常是由那些对于需要付出海森堡所提到的那种牺牲这一事实不满意的人提出的。

诚然，不满足感是科学的强大推动力量，它可能以多种方式出现。“不可能科学”导致少数特殊的不满足感的出现。它常常与人们的希望和梦想发生碰撞：关于无限能量、超光速旅行和单一实体（物体总是固定在特定的位置上）。似乎人天生就会有这样的希望，并回避使他们的希望破灭的科学。其实，不可能科学令人感到不满足是不足为奇的。最终从中受益的是科学。

注　　释

[1] 麦克斯韦给瑞利爵士的信，1870 年，麦克斯韦，《麦克斯韦的科学信件和论

文》(*The Scientific Letters and Papers of James Clerk Maxwell*),第 2 卷:1862—1873 年(剑桥:剑桥大学出版社,1995),第 583 页。

[2] 威廉·维恩,“黑体辐射与热力学第二定律的新关系”(*A New Relationship Between the Radiation from a Black Body and the Second Law of Thermodynamics*),载于《普鲁士科学院学报》(*Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*),1893 年,第 55~62 页,第 62 段。

[3] 马克思·普朗克,“论维恩辐射定律的改进”(*On the Improvement of Wien's Law of Radiation*),*Verhandl. Dtsch. Phys. Ges*. 2 (1900),第 202 页。

[4] 开尔文,“在热和光动力理论上空的 19 世纪乌云”(Nineteenth Century Clouds over the Dynamic Theory of Heat and Light),载于《巴尔的摩报告——分子动力学和光波动论》(*Baltimore Lectures on Molecular Dynamics and the Wave Theory of Light*),(伦敦:剑桥大学出版社,1904 年)。第 486~527 页。

6

19 世纪最重要的事件
麦克斯韦方程组

$$\nabla \bullet \boldsymbol{E} = 4\pi\rho$$

$$\nabla \times \boldsymbol{B} - \frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t} = \frac{4\pi}{c}\boldsymbol{J}$$

$$\nabla \times \boldsymbol{E} + \frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t} = 0$$

$$\nabla \bullet \boldsymbol{B} = 0$$

说明：完整描述了包括电磁学在内的物理现象，说明了变化的磁场如何能产生变化的电场，强调磁单极是不存在的，描述了电流和变化的电场如何能产生磁场以及电场是如何产生的。

发现者：麦克斯韦。

发现时间：19 世纪 60 年代。1884 年奥利弗·亥维赛（Oliver Heaviside）重新进行了描述。

毫无疑问，一万年以后，当我们回顾人类发展史的时候，19 世纪最重要的事件，一定是麦克斯韦发现了电动力学定律。而与这个同时代的重要科学事件相比，美国内战反而相形见绌。

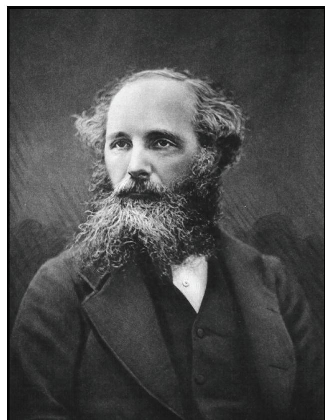
——理查德 • 费曼，《费曼物理学讲义》

费曼显然是在开玩笑，对吧？美国内战是历史上最激烈的战争之一。在这场战争中，60 多万人丧生，财产损失达 50 亿美元，400 万奴隶获得解放。自此，美国废除了奴隶制，同时也给经济、政治和社会带来了无法挽回的损失。人们至今还对这一可怕的事件记忆犹新，它怎么会在一个谦逊的苏格兰人为了描述一些几乎没有什么实际意义的奇怪效应而写出的方程面前黯然失色呢？

不过这次费曼可没有开玩笑。麦克斯韦方程组描述了一种新的现象——电磁场。电磁场完全超出了牛顿力学的范畴。麦克斯韦方程组完整地刻画了这一新现象，并且还预测了另一个不可思议的现象：存在可以穿越空间的电磁波。从这些方程出发，人们对电磁学的理解推动了从对这一现象好奇到它最终成为现代社会的基础这一过程的转变。如今，电磁学已经广泛应用于电气设备和所有基于电磁波的设备，包括收音机、雷达、电视、微波和无线通信等。在这一过程中，麦克斯韦方程对人类的影响（人类的生活方式，以及与别人、自己和世界交流的方式）大大超出了所有战争可能或业已对人类产生的影响。

麦克斯韦

麦克斯韦于 1831 年生于爱丁堡，父母在位于苏格兰西南部盖勒韦地区格伦莱尔的一幢家族庄园中将他抚养成人，并为他请了一位私人教师。10 岁的时候，家人把他送到城市里一个叫作爱丁堡学院的学校，以便接受更多的正规教育。在那里，城里的同学对乡下人打扮、口音怪异、天资平平，而且常常爱问一些天真可笑问题的麦克斯韦百般嘲笑，还给他


詹姆斯 · 克拉克 · 麦克斯韦（1831—1879）

起了个绰号“蠢货”。不过他所提出的问题并非出于无知，而确实是出于好奇。这些问题常常是：“为什么会那样？”这位年轻人后来受到威廉·汤姆逊的熏陶。汤姆逊的父母和麦克斯韦的父母是朋友，汤姆逊喜欢科学，年长麦克斯韦 7 岁。1846 年，汤姆逊已经是格拉斯哥大学的教授，研究电学。16 岁时，麦克斯韦进入爱丁堡大学学习课程。4 年后进入剑桥大学三一学院。1854 年从剑桥大学毕业后，22 岁的麦克斯韦写信给汤姆逊说他对学习电学很感兴趣，不过“所知甚少”。[1]但他学得很快，不久就赶上来了。

当时，偶尔被称作“电学研究”的这个领域只有很多人提出的一些零零散散的东西。1820 年，丹麦物理学家奥斯特（Hans Christian Ørsted，1777—1851 年）提出电流可在其周围产生磁场。不久后，法国物理学家安培（André-Marie Ampère，1775—1836 年）写出了一个方程，从数学上描述了这一现象，现称为安培定律：沿着线圈的总磁力等于通过线圈的电流。19 世纪 40 年代，麦克斯韦的导师汤姆逊（1824—1907 年）注意到了电和热的流动很相似，写出了电学方程来研究两者之间的关系。

英国科学家迈克尔·法拉第（Michael Faraday，1791—1867 年）进行了全面的研究，开展了一系列实验，于 1844 年写出了《电学的实验研究》（*Experimental Researches in Electricity*）。此外，法拉第还发现了电磁感应现象，即运动的磁体在导线中会产生电流，并且一根导线中变化的电流可在另一根导线中产生电流；以及“法拉第效应”，即在存在磁场的条件下，偏振光通过玻璃时，偏振面将发生转动，意味着磁场可以对光

产生影响。

然而，许多科学家对法拉第的工作表示怀疑。这些人采用牛顿式的眼光看待电。他们把电看成是像粒子或流体一样沿着导线流动，集中在特定的物质当中，他们认为电受到像引力一样的力的作用，这力可以跨越空间，瞬间就产生作用。对于这些科学家来说，数学对理解电磁现象至关重要。而法拉第深信以太遍布于空间，电和磁都是由以太中的应力引起的，可以机械地通过以太传播，传播速度可能是无限的。因此，法拉第深信：即便导线和导体并没有发生运动，磁和电也可以对其产生影响，产生所谓的电紧张状态（electrotonic state）。所以，光有数学还是不够的，还得要理解作用的机制。麦克斯韦后来把法拉第的观点与其他人的加以对比，写道：

> 法拉第坚持己见，认为空间遍布着力线，与数学家所认为的在一定距离下力的中心相互吸引不同，法拉第看到的是介质，而数学家看到的只是距离。法拉第从介质中的实际情形去找寻现象的来源，数学家则满足于从一定距离下施加在电流体上的动作中所作出的发现。[2]

与同时代的科学家相比，法拉第最大的问题是数学素养的缺乏。他甚至有一点惧怕数学，而更喜欢用图来表达自己的思想。例如，他把以太中的应力比作是“力线”。他的这个比喻在一定程度上是受以下事实的启发：如果把铁屑摊在一张靠近磁铁的纸上，这些铁屑就会自动排列成有规律的图案。通过电磁感应，每块小铁屑都会变成一块小磁铁，首尾相连构成光滑的曲线，从磁体的一极出发到达另一极。法拉第把这些图案看作是穿越空间的某个实际物体的表观反映。他认为，由这些线的分布、挤压和弯曲的情况就可以得出电和磁的属性。可是对这些现象，法拉第仅能从数学上进行基本的描述。同行们认为，他的实验肌肉非常充分，可惜缺少数学骨骼的支撑。

而为法拉第的实验数据提供数学支撑的就是麦克斯韦。麦克斯韦对电学研究的影响就好比是欧拉对数学的影响，他把许多看似没有联系，甚至是相互矛盾的领域统一了起来。麦克斯韦的贡献之所以大，在于他把这些领域中似乎是最独立、最富有活力的一个——光学，发展成为了一个全新领域（电磁学）的分支。欧拉是通过研究另一个领域（分析）的潜力，对数学进行了整理；而麦克斯韦则运用类比，开创了一个新的领域，对电学进行了重组。整个科学史上，可以说麦克斯韦对类比的运用已经达到了登峰造极的程度，少有人能出其右。这也导致了人类文明史上最惊人、最具决定性意义的转变的出现。

麦克斯韦的导师汤姆逊曾说："不得出物体的力学模型，我是不会满意的。有了力学模型，就能理解事物。如果得不出完全一致的模型，我就无法理解它。"[3]麦克斯韦也为力学模型所深深吸引。从剑桥大学三一学院毕业后不久，麦克斯韦曾在一个本科生俱乐部中以此为主题作了一个报告。报告基调欢快，虽思辩结构不甚清晰，不过却发人深思。[4]他告诉学生：类比说的不是相似，而是关系。科学家之所以认为类比有用，是因为自然界并不像杂志，翻开一页就知道下一页是什么。它更像是一本小说，一开始介绍的主题常常会在后面以复杂细微的形式不断出现。所以，研究一个奇怪的新现象在何种程度上类似于另一个已知的现象，再加上些必需的调整，是一种掌握未知现象的很有效的方式。

第一步：数学力

麦克斯韦在报告之前，就已经开始采用该方法对电学和磁学理论进行变换了。第一步是一篇题为"On Faraday's Lines of Force"的论文。1855 年 12 月，年仅 24 岁的麦克斯韦向剑桥哲学学会宣读了该论文。[5]论文的第一部分开头就陈述了电学的糟糕现状。有些是有实验数据支持的，而另一些则没有数据支持。有些已总结出数学表达式，有些则没有，而且表达式也不是都能很好地联系在一起。所以，研究电学的人必须要记忆

大量复杂的、不连贯的知识。这样一来，人们就很难把这些知识弄得很清楚，对此领域有所贡献。所以，很有必要对这些信息进行简化和压缩，以便更好地理解它们。麦克斯韦坦言自己不是实验学家，但他可以采用物理学上的类比，提出更适用于电学的数学表达式。请记住这些仅仅是类比。只有这样，人们才能思考得更加清楚，既不会纠缠于数学，也不会纠缠于从别处借用来的物理概念。

后来麦克斯韦提到了几个合适的类比。其中一个是法拉第提出的：电所施加的力类似于空间中弯曲的线。另一个是汤姆逊提出的，即电在空间中的流动类似于热在流体中的流动：将电荷中心类比成热源，电吸引或电排斥类比成热的流动，电势差类比成温度差等。第三个是水力学类比，即电荷就像水泵，泵出一股不可压缩的流体（如水流），泵速就相当于电荷力的强度，等等。

麦克斯韦假定法拉第的“模糊的、不很数学的”想法成立——电场是由力线组成的，力线从一个电荷出发，延伸到另一个电荷，遍布整个空间。力线上的每个点都有一定的方向和强度。麦克斯韦说，现在假定电与不可压缩流体的行为是相同的——力线相当于携带流体的微管，并且运动的阻力与速度成正比。不过针对电学的情况，还要作些修正——流体没有任何惯性。如此一来，就可以将这个由汤姆逊提出的处理流体流动的类似数学框架，应用于法拉第的关于力线的结论。麦克斯韦采用这一连续数学框架，将电磁感应、安培定律以及法拉第的其他一些物理发现，用六个定律表达了出来。在论文的第二部分中，麦克斯韦处理了法拉第的电紧张状态这一思想。他提出了一个新的变量，现称为电磁向量势（或者 A）。从汤姆逊的工作出发，麦克斯韦提出了一个包含电磁向量势的微分方程（用于描述随时间连续变化的性质）的数学结构。电磁向量势就位于这一框架中。不过他本人也承认，这一结构并不能“解释一切”，它“甚至缺少真正的物理理论的影子”，似乎并没有提出任何新的东西。不过该结构的确为法拉第的研究提供了“数学基础”。这也是任

何物理理论最终形成的必要条件。[6]

法拉第接到麦克斯韦论文时的第一反应是几乎被“这样一个数学上强有力”的应用“吓住了”。可随后他又很高兴，因为这一努力获得了成功。[7]

第二步：重要的类比

麦克斯韦工作的第二步是于 1861 到 1862 年间写成的一篇叫做《论物理力线》（*On Physical Lines of Force*）的论文。这篇论文堪称科学史上最伟大的应用类比的例子。麦克斯韦先是说明了自己的意图，即“以机械的观点检视磁学现象”，然后参考了汤姆逊在理解法拉第效应时所采用的类比：如果磁场可以使光的偏振面发生移动，那么磁力线上的各个点就好比都是不断旋转的微小“分子涡旋”。它们把转动的一部分传递给流经的光波。

之后，麦克斯韦进一步扩充了这一设想。我们干脆说磁场是由这些旋转的“小室”构成的。这些小室的轴沿着磁力线的方向，就像是穿在磁力线上一样。磁场的强度越大，小室旋转的速度越快。不过，麦克斯韦很清楚，从原理上说相邻线上的小室的旋转方向不可能相同——比如说都沿着顺时针方向。如果旋转方向相同，那么相邻两根线上的小室就会相互摩擦，而这是错误的。为了使设想成立，他又假设小室之间的空间是由一种与工程师们常说的“空转轮”类似的物体填充的——小珠子。这些小珠子与小室相接触，沿着逆时针方向旋转，从而允许小室沿着顺时针方向旋转。当相邻单元以相同速度旋转时，这些小珠子保持不动。但是涡旋速度的变化会使小珠成行移动，从一个单元进入到另一个单元中去。由此，麦克斯韦认为这些小珠子的作用与电流是相同的。

该模型显示出了由介质的机械运动产生的电磁效应——变化的磁场产生电流，变化的电流也可以产生磁场。以太中的推挤可产生所有电学和磁学效应。这一点法拉第和其他人已经注意到了。该模型甚至还产生

了法拉第电紧张状态的力学概念，即在有磁场而没有电流的情况下会发生什么现象。电紧张状态类似于空转轮转动时（没有运动）的冲力。

麦克斯韦的设想

麦克斯韦的模型

1861 年春天，麦克斯韦将这一思想写进论文，在 3 月到 5 月间连载。之后，他就去往格伦莱尔避暑度假了。他并不奢望自己能提出电磁学的设想或者表述。他想说的无非是这个奇怪的模型与其他电学和磁学现象并无二异，因此模型的数学表达也适用于后者。麦克斯韦特别提到，他的模型就好比是一架天象仪，或者是人们常常在自然历史博物馆看到的太阳系模型。在太阳系模型中，行星就是穿在杆上的球，沿着中间的球（也就是太阳）作机械摆动。组装这样一个模型（把所有知道的东西都融入进去）的意义在于一旦模型完成，人们就可以去研究太阳系作为一个整体是如何运转的，而且常常可以看到仅从部分来看所看不到的东西。

度假期间，麦克斯韦注意到他在模型中遗漏了一些重要的东西。像所有固体一样，前面提到的“小室”至少也要有一点弹性。但是弹性会在模型中产生一定的效应，而这一点他并没有考虑。小室挤压珠子时，

珠子无法发生移动（比如在绝缘材料中）。但因为小室有弹性，它们就会对珠子产生一点推力，直到运动被材料中的力相互抵消，就像橡胶球受到固定不动的力的作用一样。如果把力撤掉，小室和珠子都会恢复原来的形状。麦克斯韦把这称为“电位移”，电位移的大小由电动力的强度和物体的性质决定。同时，他认识到了必须把这一点考虑到数学表达中去。在该过程中还要考虑安培定律的一个小型修正因子。

更具革命性的观点是：任何弹性物体都可以以波的形式从一处向另一处传播能量。麦克斯韦已经表明，以太（电磁现象的介质）必须至少得有一点弹性。介质可通过交互前进且相互垂直的电学和磁学效应，将能量以波的形式从一部分传输到另一部分，从空转轮到小室，再从小室到空转轮，如此周而复始。与光波一样，这些波也可以发生反射、折射、产生干扰和偏振等现象。麦克斯韦由此出发，并假定横向振动是由纯粹的机械力传播的，发现了横波在以太中传播的速率。以几年前测定了电学常数的德国物理学家鲁道夫·科尔劳施（Rudolf Kohlrausch）和威廉·韦伯（Wilhelm Weber）的工作为基础，麦克斯韦计算出速率为 310 740 千米/秒，亦即 193 088 英里/秒。但是 12 年前，阿曼德·斐索（Armand Fizeau）所测出的光速为 314 858 千米/秒，亦即 196 647 英里/秒，与麦克斯韦的计算值很接近。于是麦克斯韦写道：“在我们假想的介质里，根据科尔劳施和韦伯的电磁实验计算得出的横波速度与通过斐索的光学实验计算得出的光速完全吻合，因此顺理成章得出推论‘产生电磁现象的介质的横向振动是由光组成的’。”[8]

麦克斯韦将模型中这两个具有革命性意义的新特征作为他论文的第三部分，于 1862 年发表。

保住婴儿

两年后，麦克斯韦在一篇题为《电磁场的动力学理论》的论文中迈出了关键性的一步。论文写于 1864 年末期，发表于 1865 年初。在论文

中，麦克斯韦引用了早期的机械类比，而实际上却只是要摒弃它，目的是将所有结果——包括位移电流和光是电磁波的思想等，以一组独立方程的形式呈现出来。“这样一来，我们就得出了一个复杂机制的概念。这一概念涉及各种运动，但同时因为这些部分是相互联系的，根据一定的关系，某一部分的运动是与其他部分的运动联系在一起的。这些运动之间的相互影响通过连接部分的相对位移所产生的力来实现，采取的是弹性的形式。这样一个机制必须受到动力学一般性定律的约束。我们应该知道各部分的运动之间的联系的形式。”[9]在后边的几段中，麦克斯韦接着写道：“为了用符号计算处理这些结果，我需要将结果用电磁场通用方程的形式表达出来。”他接着列出了分属 8 个大类的 20 个方程。[10]

这是科学史上最著名的采用类比的例子。麦克斯韦的成就本身就是用著名的类比形式表达的——“麦克斯韦倒掉了洗澡水，保住了婴儿”——不过对他来说，这里的洗澡水乃是婴儿之父。

论述

1873 年，麦克斯韦发表了《电磁通论》(*A Treatise on Electricity and Magnetism*)，完整地展示了他用惊人的类比所建立的科学分支。在随后至少 10 年的时间里，人们都必须要学习这一类比形式。该书有 1000 多页，书的内容很难理解和消化。麦克斯韦在写书之时，并没有考虑读者接受能力，因此未对书中内容进行提炼和简化。他写书的目的是求全而不是经济利益。例如，在关键的一章“电磁场的一般方程”中，他分 12 步对自己的工作加以总结，分别标成 A 到 L，每一步中都有一个或一组方程。他写道：“这可以看作是我们所考虑的量之间的主要联系。”有些是可以合并到一起的。“不过，我们的目的并不在得到简洁的数学公式。”而且，这些公式对那些只对实际应用感兴趣的人来说，使用起来是极其困难的，尤其是矢量势 $\boldsymbol{A}$ 和标量势 ψ。

历史学家也对麦克斯韦的《电磁通论》感到迷惑。他在书中（以及其他地方）并没有告诉读者电磁波是如何产生和发现的。电磁波的思想是麦克斯韦一生的工作中最令人兴奋，而又最出人意料的一个。他的这一反常做法就好比一位天文学家的研究工作预言出一颗新的行星的存在，但却不去想如何找到一个望远镜去观察它，或者告诉别人怎样去观察。他的这一做法委实令人费解，不解释似乎说不过去。有些历史学家说麦克斯韦对电磁波的兴趣并不及光和以太那么大；另一些人则说他想不出什么办法，能产生出电磁波，并对其进行探测。还有一些人干脆说他没时间去想这一问题。这些解释都不够令人信服，虽然麦克斯韦的工作量在《电磁通论》出版之时急剧增加。1871 年，麦克斯韦受命监督英国剑桥大学新卡文迪许实验室的建设。1874 年，他又接手了对亨利 • 卡文迪许（Henry Cavendish）的论文进行编辑的任务。另外，他还是《大英百科全书》（*Encyclopedia Britannica*）第九版的科学编辑。因为这些项目的原因，麦克斯韦做研究的时间就很少了。

但是，麦克斯韦对“以太的巨大海洋”能否探测到这一问题仍抱有兴趣。以太是看不见的，人们对它也所知甚少。麦克斯韦在《大英百科全书》中的“以太”这一部分中写道：人们甚至不知道像地球这样的致密天体是不是像鱼儿在水里游动那样在以太中运行的，并且在运动时还拖拽着一些以太；人们也不知道以太是不是“像海水经过船拖拽的网上的网眼一样”经过天体的。他曾优美地，略微激动地写道：

> 空间中没有路标，空间中的任何部分与其他部分都是类似的，所以无法说出路标的位置。我们位于平静的海面上，没有星星、指南针、大气探测、风或潮汐，也不知道前进的方向。我们找不到可以放出去帮助定位的圆木。我们可以计算出相对于附近天体移动的速度，但却不知道这些天体在空间中如何运动。[11]

幸好在介质速度不同时，波在介质中传播的速度也不同，由此麦克斯韦找到了一个发现波的技巧。例如，声波总是以相同的速度传播，在空气中的速度大约是每秒 1 100 英尺。它的传播速度仅仅取决于传播介质（空气分子）的性质。如果有风，风就会带着声波运动。而声波在空气中传播的速度不变，所以在地面上的人看来，声速就变大或变小了。因此，有风的时候，声波在不同方向上的传播速度就是不同的。

光波与声波类似。地球在绕太阳运行时，会“带走”一些以太，因此它相对于以太的运动速度会发生变化，这时会产生以太风或者以太漂移。于是，不同方向上的光速就是不同的。这种与静止以太中的光速之间的差异是微小的，只有一亿分之一。那么可以测量到这种差异吗？

在地球上恐怕测量不到。麦克斯韦写道：如果实验者向各个方向发射光束，并且令光束返回，那么传播时间上亿分之一的不同是“很难察觉的”。“唯一可行的办法是从地球上接近于黄道相对点的位置观察木星，通过观察到的木星卫星的偏食加以推断，得到光速值，然后再对这些值进行比较。”1879 年 3 月，麦克斯韦联系了英国剑桥航海天文历编制局的主任，向他询问有没有这方面的已有研究。他带着惯常的谦逊在信中写道，“我不是天文学家，但现在我所知道的唯一的”以太漂移的测量办法是对木星卫星的偏食的明显推迟进行精确的测量。[12]

此时，麦克斯韦已经有了腹部癌症的症状。同年 11 月，麦克斯韦去世，这位创造力惊人的科学家年仅 48 岁。他静静地执着于自己的追求，默默改变着世界。费曼称赞他对世界的影响超出了美国内战。

麦克斯韦没有完成自己的事业就离开了。因为种种原因，他的工作所预示的激动人心的思想是他本人所没有研究过的。第一个是产生和探测电磁波的问题；第二个是测量以太漂移的问题；第三个就是用简洁的方式对方程进行重写，方便实际应用的问题。随着电报的发展，最后一条变得日益重要起来。这三个问题在麦克斯韦死后的十余年里一一得以解决。

海因里希·赫兹和电磁波的发现

海因里希·赫兹（Heinrich Hertz，1857—1894 年）在汉堡出生和长大。1878 年在柏林，赫兹在赫尔曼·冯·亥姆霍兹（Hermann von Helmholtz）的指导下开始学习。亥姆霍兹当时正在研究麦克斯韦的电动力学。他劝说这位 22 岁的聪明年轻人去研究一个亥姆霍兹自己所提出的实验问题。这一问题如能解决，就能确认麦克斯韦理论的某个特征，问题解决者也将被授予奖项。但赫兹拒绝了，他担心几年的时间耗进去之后却得不出什么确定性的重要结果。因此赫兹主要还是集中精力去完成博士论文。1885 年，赫兹来到卡尔斯鲁厄，进入了设备精良的实验室，他利用良好的实验条件创造性地开展工作。1886 年，他偶然发现震荡电流可使火花跨过与附近线圈之间的微小间隙。赫兹以此写出一篇题为“空气中的电磁波及其思考”（On the Electromagnetic Waves in Air and Their Reflections）的论文，发表在 1888 年 7 月的《物理学年鉴》上。赫兹测出了电磁波的波长，表明电磁波有着与其他波一样的性质——包括可以反射、折射、干扰、被极化，以及具有有限的速度等——极好地验证了麦克斯韦的理论。

同时，英国利物浦的物理学教授奥利弗·洛奇（Oliver Lodge）注意到了震荡电流在导线中所产生的波。1888 年 7 月，洛奇用自己的实验结果完成了一篇论文，然后踏上了去往阿尔卑斯山的火车作远足旅行。途中，他拿出自己的阅读材料——当月的《物理学年鉴》，并了解到了赫兹的工作。他极为沮丧，自己本想参加 9 月在巴思举行的英国科学促进会的年度会议，庆祝自己的发现。可现在，洛奇发现赫兹的工作使自己的研究成果黯然失色。不过，洛奇却为赫兹实验的完美而欢欣鼓舞。赫兹的实验更加全面——他不仅在导线中，而且在空气中也探测到了电磁波。

巴思会议上，赫兹首次向科学界公开介绍自己的发现，当时的情形非常令人激动。[13]由于数学和物理分部的原主席身体不好，因此由爱尔

兰物理学家乔治·菲兹杰拉德（George FitzGerald，1851—1901年）接替。菲兹杰拉德研究产生电磁波的可能性将近有10年，所以他早已准备好隆重介绍赫兹工作的重要意义。会议热闹非凡，有托马斯·爱迪生（Thomas Edison）发明的新式蜡模留声机以及乔治·萧伯纳（George Bernard Shaw）所做的关于“社会民主”的演讲。可会议的风头却被菲兹杰拉德带来的消息抢了：电磁力不是在一定距离下发挥作用的，而是通过在以太中传播的电磁波发挥作用的。他宣布：“1888年将载入史册，因为赫兹在德国用实验解决了这一伟大问题。”《时代》杂志听到菲兹杰拉德的讲话后，称该消息是“划时代的”。然而，麦克斯韦思想的确认过程也使长久以来一直存在的对麦克斯韦公式的不切实际而产生的不满浮出了水面，菲兹杰拉德提到了与会者要“谋杀 ψ”或者起码要修正矢量势 $\boldsymbol{A}$ 的尝试，大家的一致意见是要取消某些概念。

电磁波的产生和探测这一惊人消息——麦克斯韦的工作暗示出了这一点，不过他本人并未指出——也为方程本身出人意料的作用提供了一个经典阐述。就像赫兹有一次提到麦克斯韦方程所说的那样：“人们摆脱不了这种感觉，即这些数学公式有其独立的存在和智能，它们比人更聪明，甚至比它们的发现者还要聪明。人们从方程中所得到的，比最开始所投入的还要多。”[14]

阿尔伯特·迈克尔逊和以太的不存在性

麦克斯韦关于以太漂移的信送到了航海天文历编制局的主任手中。1880年1月初，也就是麦克斯韦去世后两个月，皇家学会读到了这封信。不久，它就在《自然》杂志上发表了。有位读者对此非常着迷，他就是美国物理学家阿尔伯特·迈克尔逊（Albert A. Michelson，1852—1931年）。迈克尔逊研究生毕业于马里兰州安纳波利斯的美国海军学院。毕业后留在该学院教授科学。他深深沉迷于光速的测量问题，曾于1878年和1879年两次逃开7月4日的学院传统庆典，不间断地从事研究工作。在1879

年的测量中，迈克尔逊将一束光打到 200 英尺之外，再反射回来，获得了前所未有的精度，为这位 27 岁的年轻人在科学家中间赢得了声誉。8 月 29 日出版的《纽约时报》也在头版登出了这一新闻。可惜，迈克尔逊的名声并没能打动学院，没有给他免除规定的海外勤务。不过迈克尔逊托人给自己请好了假，这样他就能在 1880 年初去欧洲亥姆霍兹的实验室学习物理学。麦克斯韦死后，他在读到了 1880 年 1 月出版的《自然》杂志上发表的信件之后，发明了一个装置，称为“干扰折射计”。干扰折射计利用半透镜，利用光线折射（弯曲）将一束光分成两束，然后将这两束光相互垂直地发射出去，再反射回来。如果使这两束光发生干扰，那么它们在以太中沿着不同方向传播时所产生的相差就与波长的一部分相当了。但是这种微小的差别是“很容易就能测出的”，迈克尔逊在写给《自然》杂志的信中写道。[15]他还向自己的孩子们解释实验计划，请他们想象一下“两个比赛的游泳者，一个向河的上游游动，然后再回来；另一个也游相同的距离，但他是向着河对岸游，然后再游回来”。“关键的一点是，只要河水是流动的，那么后一个游泳者总是会赢。”[16]

1881 年的首次实验没有探测到任何以太漂移，实验设计上似乎也有漏洞。之后，迈克尔逊退役，来到俄亥俄州克利夫兰的凯斯应用科学学院，与另一位实验者爱德华·莫雷（Edward Morley，1838—1923 年）合作，对仪器进行扩充和修正。这次实验的灵敏度相当惊人，达到了四十亿分之一，不过仍然没有探测到任何漂移现象。迈克尔逊对这种没有结果的实验感到厌倦，甚至失去了信心。他和莫雷放弃了继续测量的计划。但其他科学家，包括乔治·菲兹杰拉德、荷兰物理学家亨得里克·洛仑兹（Hendrik Lorentz）和法国数学家亨利·庞加莱（Henri Poincaré）在内，都开始义无反顾地尝试证明迈克尔逊–莫雷实验和以太的存在性。他们的努力为爱因斯坦发现狭义相对论打下了基础。1907 年，因为在完成这一重要实验中所发挥的作用——受麦克斯韦信件的启发——迈克尔逊成为第 10 位诺贝尔物理学奖获得者。他也是获此殊荣的第一位美国公民。

奥利弗·海维赛德和“麦克斯韦方程组”

标准化的麦克斯韦方程组很大程度上是由奥利弗·海维赛德（Oliver Heaviside，1850—1925 年）完成的。海维赛德是一位电气工程师。他自学成才，性格古怪，行为乖张。他发现了被称为海维赛德层的电离层。他常被人们称作是“现代科学中的最后一位业余科学家”。[17]他 16 岁离家，从未在大学工作过，一生贫困，靠亲戚、朋友救济，领政府养老金。他唯一的工作是当了 4 年的报务员。他对改进能量沿着电报线的流动等实际问题非常感兴趣。海维赛德大部分当代的数学知识都是通过自学习得，并以此作为改善电磁理论现状的全新手段。例如，他把虚数引入了电学中。当看到麦克斯韦的《电磁通论》一书时，他的反应和麦克斯韦本人看到当时的电学状况多少有些类似：对实用派来说过于复杂了，读者需要同时具备太多的知识。麦克斯韦的公式与当时人们对电报的日益关注并不相宜。这一公式建立在向量势 $\boldsymbol{A}$ 和静电势 ψ 的基础上，尚存在“一定距离下的作用”这一观点的残迹。而电报所涉及的则是电磁能沿着特定线路的流动。

19 世纪 80 年代，实用科技的需求的确在很大程度上推动了电磁学的发展。[18]当时许多电磁学的研究者都提出过巧妙的物理模型（模型中有轮子和连接带），来设想在麦克斯韦理论中，电能是如何从一个地方流动到另一个地方去的。许多人对麦克斯韦采用向量势 $\boldsymbol{A}$ 和静电势 ψ 都感到特别困惑。

1883 年，海维赛德在《电工》（*Electrician*）杂志上发表了一系列文章，开始研究如何对麦克斯韦的工作加以改写，以便适应电报线和电路中电的流动研究这一实际背景。海维赛德后来写道：“只有改变表述的形式，我才能看得更清楚。”[19]他利用业余时间自学的知识足以够用，可以既不受当代数学知识的局限，也不受盛行的物理观点的影响。他的观点是实用化的，对他来说，重要的是各个点上的能量以及计算能量是如何

沿着导线的路径流动的。他倾向于用简单直接的方法如魔法般地将这一观点表达出来，就像海维赛德在一篇科学论文中某段的起始句中所写的那样："能量从一个地方流动到另一个地方的时候，所穿越的是中间空间。"[20]之后，他大胆改写了麦克斯韦方程组，分别用 $\boldsymbol{E}$ 和 $\boldsymbol{H}$ 表示各状态下的电力和磁力，用 $\boldsymbol{D}$ 和 $\boldsymbol{B}$ 表示电流和磁流。最后，他把麦克斯韦方程组彻底缩写，变为 4 个方程。这 4 个方程有着令人愉悦的对称美——两个是电学方程，两个是磁学方程，很明显是并列的。麦克斯韦方程组的改写如此之彻底，以致人们有时会将新方程称为"海维赛德方程组"。[21]在自由空间中，方程的形式如下：

$$\begin{aligned} &\operatorname{div} \varepsilon \boldsymbol{H} = \rho \qquad && \operatorname{curl} \boldsymbol{H} = k\boldsymbol{E} + \varepsilon \dot{\boldsymbol{E}} \\ &\operatorname{div} \mu \boldsymbol{H} = 0 \qquad && -\operatorname{curl} \boldsymbol{E} = \mu \boldsymbol{H} \end{aligned}$$

在电荷存在的情况下，更复杂的形式是：

$$\begin{aligned} &\operatorname{div} \varepsilon \boldsymbol{H} = \rho \qquad && \operatorname{curl} (\boldsymbol{H} - \boldsymbol{h}_0 - \boldsymbol{h}) = k\boldsymbol{E} + \varepsilon \dot{\boldsymbol{E}} + \boldsymbol{u}\rho \\ &\operatorname{div} \mu \boldsymbol{H} = \sigma \qquad && -\operatorname{curl} (\boldsymbol{E} - \boldsymbol{e}_0 - \boldsymbol{e}) = g\boldsymbol{H} + \mu \dot{\boldsymbol{H}} + \boldsymbol{u}\sigma \end{aligned}$$

海维赛德谦虚地把他的 4 个方程称作"换了新装的麦克斯韦方程组"。[22]其实他也曾热情善辩地宣传这些方程，认为它们比麦克斯韦的或者其他人改写的版本都要好。例如，1888 年巴思会议后不久，海维赛德就写了一个简短的评论，毫不留情地攻击在传播方程中还在使用电势 ψ 和向量势 $\boldsymbol{A}$ 的做法，并认为这种做法是"形而上学的"（这个词在科学家看来是侮辱性的），是"数学小说"。[23]毕竟人们测量的是电力 $\boldsymbol{E}$ 和磁力 $\boldsymbol{H}$，而不是势。$\boldsymbol{E}$ 和 $\boldsymbol{H}$ 提供的是场的实际状态信息；电流流动时，两者将随之传播。保留 ψ 和 $\boldsymbol{A}$ 将使人陷入"无法穿越的势的迷雾之中"，甚至会产生前后不一致。回顾巴思会议，海维赛德主张"弃用 ψ 和 $\boldsymbol{A}$"。他得出结论说："只要把 $\boldsymbol{E}$ 和 $\boldsymbol{H}$ 看作变量"，那么麦克斯韦的理论就是可行的。

海维赛德版本的麦克斯韦方程组很快就被包括赫兹在内的著名电磁学研究者所采用。到 19 世纪 90 年代，整个科学界也由此焕然一新。自此之后，麦克斯韦方程就几乎再未改变过，本章开头的版本就是取自 J.D. 杰克逊（J. D. Jackson）所著的标准教科书《经典电动力学》（*Classical Electrodynamics*）。

海维赛德改写麦克斯韦方程组所取得的成就其实是通过一次类比获得的。在回顾了海维赛德的工作之后，菲茨杰拉德把麦克斯韦比喻为一个将军。这位将军占领了新的疆土，却没有时间去找出最好的道路，绘制出系统的地图。“这件事正好就留给赫维赛德，”菲茨杰拉德写道，“麦克斯韦的《电磁通论》一书因为战斗中良好的阵形和挖好的营地等，妨碍了内容的表达。海维赛德则对道路进行了清理，开启了一条直通路线，建成一条宽阔的路，并探索了国家的大片领土。”[24]

然而，麦克斯韦是一位在特殊地形上战斗的奇怪的将军。正如费曼所说：麦克斯韦征服的疆土比他的战斗本身更广泛、更有说服力，它对人性所产生的影响胜过任何将军。



茶　歇

克服病觉缺失，恢复人文活力

西蒙·沙玛（Simon Schama）所著的《不列颠史》记载了英国悠久的历史。该书有 1500 页之多，一部几集的纪录片就是以此为题材拍摄的。可这本书却丝毫没有提及麦克斯韦，更没有提在 20 世纪电子革命中，麦克斯韦在电气化、光、热和通信的开创性工作中所作出的世界性贡献。作者还略去了有关英国科学家和工程师科学贡献的所有出处。而正是这些人变革了英国和整个世界。

的确，对科学的忽视已经成为史书的通病。而尤为让人恼火的是，一些打着服务大众、服务贫困、受压迫的人们旗号的书也是如此。比如，1980 年首次出版的霍华德・津恩（Howard Zinn）的著作《美国人民史》（*A People's History of the United States*）已经卖出了不止一百万本，已经成为美国最具影响的历史著作之一。作为一本在中小学和大学里广为人知的教科书，该书称将把重点放在“那些被隐藏起来的历史上的事件（哪怕只是短短的一瞬）。在这些事件中，人们表明自己有能力去反抗、团结起来，赢得胜利。”

然而，津恩却没有提到人们在研究科学过程中的抵抗、团结和胜利。比如，该书就没有提到人们在降低儿童死亡率、延长人均寿命或者发明公共交通运输系统方面所作出的努力；也没有提到因为领导了“绿色革命”而于 1970 年获得诺贝尔和平奖的诺曼・布劳格（Norman Borlaug）。诺曼还帮助解决了几百万人的吃饭问题。还有一个被忽略的人是微生物学家莫里斯・希勒曼（Maurice Hilleman）。他发明的疫苗所拯救的生命比在人类全部战争中丧生的人还要多。而津恩却为这些战争着墨颇多，用了几章的篇幅。

津恩在书中也没有提到大规模电气化，尽管他在说明为避免社会底层人们暴动，“应给予他们足够帮助”时讨论了电的单位成本问题。同样没有被提到的还有蒸汽机和内燃机，虽然他在说明有关种族隔离、工会、罢工和剥削美国印第安人时还是讨论了铁路的作用。

简而言之，津恩认为科学上的变革对“人民”是不重要的。在他看来，历史就是意识形态的一场大型表演。如果说科学在这场表演中也能起到一点重要作用的话，只可能是用它能制造武器。然后不同意识形态的党羽们用这些武器相互厮杀。

对科学贡献的省略并不会使这些史书看上去有什么瑕疵。如津恩指出的那样，史学家不可避免地会去选择和强调某些事实，尽管他们有责任尽量不去宣扬意识形态上的倾向（不管是有意还是无意的）。不过，从

“人民”的立场上看，津恩的做法的确使《美国人民史》一书有了瑕疵。对一度流行的可怕疾病（如脊髓灰质炎和脑炎）的成功征服从根本上改变了人类对生死的看法。天文学的发展和进化论的发现影响了人们的时空观，以及人类对自身在自然界中的地位的反思。这些事件都是在《美国人民史》一书所讨论的时间框架内发生的。有些科学进展虽然不是由美国人做出的，但它们也同样极大影响了人们去寻找“知道了什么、应该做什么和期望什么”这些问题的答案的途径。

忽视科学影响的不光只有沙玛和津恩。许多当代的科幻作者在书中所写的角色不过是看似不被技术训练和技术装置影响的老小孩儿。有些作者，像乔纳森·弗兰岑（Jonathan Franzen）、伊恩·麦克尤恩（Ian McEwan）、尼尔·斯蒂芬森（Neal Stephenson）和大卫·福斯特·华莱士（David Foster Wallace）有所不同，他们在书中描写了对技术感兴趣，同时也受技术影响的角色。不过这些人可能会遭到评论人的激烈批评。

比如，约翰·班维尔（John Banwille）在评论麦克尤恩的《星期六》（*Saturday*）一书时，嫌作者“刻意显示自己的技术知识”，还抱怨“书中太多大词”。这其实并不难理解，经过训练成为技术专家的人不光会习惯于用大词，甚至他们的说话、做事的方式也会随之受到影响。技术上有竞争力的人常常会以自己掌握的技术为乐，在行事时也会对其加以运用。这也是麦克尤恩能够如此娴熟地把握住角色的原因。

忽视科学技术对现代生活的影响与“两种文化”并无关联。相反，这表明有些作家和学者的著作中存在盲点。这些人对周围的世界本应有清晰的认识。就严重性来说，忽视科学技术甚至甚于健忘。对于此种状态，我们可以用 McEwan《星期六》一书中的角色使用的一个“大词”来描述。这个词就是“病觉缺失”（anosognosia）。它是一个医学词汇，是希腊词 agnosia（即“不了解”）和 nosos（即疾病）二者的组合，意思是对自己的病态缺少清晰的认识，也就是说不知道自己病了。

那么，导致“病觉缺失”的原因是什么？我认为有四个主要因素。

第一就是戏剧性：科学技术上的变革不像历史上的转折性事件那样有着激动人心的场景。预示科技变革的不是血染的战场或者大人物之间的对抗。科技变革是以另一种方式展示出来的，要把它与战场等之间的不同以戏剧的形式体现出来非常困难。第二，即便是某些所谓的进步学者也还抱着一种马克思主义式的希望，就是人们可以彻底改造自己、重塑世界，通过一次革命性的爆发获得自由。而如果承认人们要依赖于科学和技术的话，无疑会令这一希望黯然失色。第三是人们对要经过额外训练才能习得的专业知识的畏惧。

最后，也是最重要的，人文领域的学者往往把自己看的太重要了。他们以为自己提出的问题都是重要问题，可以帮助人们去应对世上的危险。但如果“人民”的命运和科学技术之间的联系与和意识形态（或者说压迫人民的那些人）之间的联系同样紧密的话，意识形态的主导地位就会被削弱，或者至少会打折扣。对上面所说的学者们来说，这似乎非常要命。所以最好还是集中精力，好好研究人性的特点，而不是“什么是可能的”！这也是为什么会有如此之多的人文项目虽然无懈可击，却显得死气沉沉的原因所在。而且，这种所谓的集中精力其实是伪装了的利己主义。于是，意识形态本身只是没有经验支撑的信仰架构而已。

克服病觉缺失就要承认真正的人性不像我们所希望的那么激动人心，要求我们不要再幻想存在实现自由的捷径，要求我们承认人性的重要问题是要通过许多不同的学科才能得以解决的。承认上述几点不仅不会威胁到人性，反而会使它得到加强。只有清楚了科学所揭示的世界的层面和可能性，并将人性的探究与人的层面和可能性联系在一起后，人性才能最有效地给出“知道了什么、应该做什么和期望什么”这些问题的答案。

注　释

[1] P. M. Harman 编，《麦克斯韦的科学信件和论文》（*The Scientific Letters and*

Papers of James Clerk Maxwell)，第 1 卷，剑桥：剑桥大学出版社，1990 年，第 254 页。

[2] 麦克斯韦，《电磁通论》(*A Treatise on Electricity and Magnetism*，纽约：Dover，1954 年)，第 ix 页。

[3] 汤姆逊，《开尔文在巴尔的摩的讲义与现代理论物理学》(*Kelvin's Baltimore Lectures and Modern Theoretical Physics*)，R. H. Kargon 和 P. Achinstein 编，剑桥：麻省理工学院出版社，1990 年，第 206 页。

[4] 麦克斯韦，“关于‘自然中类比’的使徒随笔”(*Essay for the Apostles on 'Analogies in Nature'*)，载于《麦克斯韦的科学信件和论文》，第 1 卷，P. M. Harman，剑桥：剑桥大学出版社，1990 年，第 376~383 页。

[5] “On Faraday's Lines of Force”，载于《麦克斯韦的科学论文》(*The Scientific Letters of James Clerk Maxwell*)，第 1 卷，W. D. Niven，纽约：Dover，1965 年，第 155~229 页。

[6] 麦克斯韦，《科学论文》，第 207 页。

[7] 见 1857 年 3 月 25 日法拉第致麦克斯韦的信，载于麦克斯韦的《麦克斯韦的科学信件和论文》，第 548 页。

[8] “论物理学力线”(On Physical Lines of Force)，载于《科学论文》，第 500 页。

[9] 麦克斯韦，《科学论文》，第 533 页。

[10] 1868 年，麦克斯韦写了一篇很短的论文，“A Note on the Electromagnetic Theory of Light”，载于《科学论文 II》，第 137~143 页。在论文中，麦克斯韦承认在自己之前的电磁现象方面的工作中，与光的联系本身是不容易理解的。他用“最简单的形式”——4 个定理，重新叙述了这一联系。不过这 4 个定理还不是“麦克斯韦方程”。

[11] 引自 Dorothy M. Livingston，《光学大师》(*The Master of Light*，纽约：斯科里布纳，1973 年)，第 100 页。

[12] 麦克斯韦，“On a Possible Mode of Detecting a Motion of the Solar System through the Luminiferous Ether”，《自然》，1880 年 1 月 29 日，第 314~315 页。

[13] 会议的详情可参见 B. J. Hunt，*The Maxwellians*，伊萨卡：康奈尔大学出版社，1991 年，第 7 章。

[14] 引自 E. T. Bell，《数学大师》(*Men of Mathematics*，纽约：西蒙与舒斯特公司，1937 年)，第 16 页。

[15] 迈克尔逊，“The Relative Motion of the Earth and the Luminiferous Ether”，《美国科学期刊》(*American Journal of Science*，22，1881)，第 120 页。

[16] 引自 Livingston，《光学大师》，第 77 页。

[17] D.S.L. Cardwell，《英格兰的科学组织》（*The Organization of Science in England*，伦敦：海尼曼出版社），第 124n 页。

[18] 通量理论具有特殊的重要意义。"当时有关麦克斯韦理论的研究工作很容易偏离主题，变成纯粹的数学解释。能流原理的发现使人们把注意力牢牢集中到场的物理状态上来。"Hunt，*The Maxwellians*，第 109 页。

[19] Oliver Heaviside，《电磁理论》（*Electromagnetic Theory*），第 1 卷，纽约：切尔西，1971 年，第 vii 页。

[20] Oliver Heaviside，《电学论文》（*Electric Papers*），第 2 卷，纽约：切尔西，1970 年，第 525 页。

[21] Hunt，*The Maxwellians*，第 122 页。

[22] 引自 Hunt 的 *The Maxwellians* 一书的附录，"From Maxwell's Equations to 'Maxwell's Equations'"，第 247 页。

[23] Heaviside，"论势传播的形而上学实质"（*On the Metaphysical Nature of the Propagation of Potentials*），《电学论文》，第 2 卷，第 483~485 页。

[24] Hunt，*The Maxwellians*，第 128 页。



方程中的“名流”
爱因斯坦质能方程

$$E = mc^2$$

解释：能量和质量可以相互转化。能量等于质量乘以光速的平方。
发现者：爱因斯坦。
发现时间：1905 年。

不久前，我在一本电影杂志上看到了女演员卡梅隆·迪亚兹（Cameron Diaz）的一个采访。快要结束的时候，采访者问她是否有很想知道的事情。迪亚兹说很想知道 $E = mc^2$ 到底是什么意思。说罢，两个人都笑了，迪亚兹说她是真的很想知道。采访就此结束。

——戴维·波丹尼斯（David Bodanis），《$E = mc^2$：世界上最著名的方程》（$E=mc^2$: *A Biography of the World's Most Famous Equation*）

$E = mc^2$ 是有史以来最著名的方程。这个方程曾作为《时代》杂志的封面图案；它还是一部“传记”的主角，享受和人一样的待遇；海莉·弗兰纳甘（Hallie Flanagan）有一部戏就是用它做的标题，该剧曾在大萧条时期风靡联邦歌剧院。在诗歌和流行音乐中，它也备受青睐。上了年纪的人应该还记得当年的热门单曲《爱因斯坦冲啊冲》

(*Einstein A Go-Go*)。这首歌由 20 世纪 80 年代的电子流行乐队 Landscape 演唱，歌词有一句是："你最好小心些，你最好谨慎些，因为阿尔伯特说 E 等于 m 乘以 c 的平方。"最近，歌手玛利亚·凯莉（Mariah Carey）推出了她的新专辑，专辑的名字就是《$E = MC^2$》，借此说明自己的创作初衷。在上世纪 90 年代所谓的科学大战中，法国女权主义哲学家露丝·伊利格瑞（Luce Irigaray）曾断言 $E=mc^2$ 是个有性别的方程，因为方程中的光速被平方了，这一说法引起了轩然大波。[1]在世界各地的邮票上、电影中（《摇滚校园》(*School of Rock*)）、通俗科幻小说里（丹·布朗的《天使与魔鬼》(*Angels & Demons*)）以及各种卡通和电子游戏中，也常能看到该方程的身影。

物理学家斯蒂芬·霍金（Stephen Hawking）在写作一本面向普通读者的书时，曾被人警告不要在书中使用任何方程。因为（或许人们就是这么告诉他的）每多写一个方程，销量就会减半。因此，霍金决定在《时间简史》里一个方程也不写。但是在书中还是有几处提到了方程 $E=mc^2$。不过，该书的销量不仅没有因此受到任何影响，而且还成为有史以来面向普通读者的最畅销科学书籍之一。

所有这一切都增加了我们的好奇心：$E=mc^2$ 是不是已经超越了方程的层面，而变成一介"名流"了呢？名流就是那些人们听说过，却又不了解的人。同样的道理，人们都知道该方程，也知道其意义重大，但却说不出个所以然来。关于它的说法有很多，可是人们还总是感觉自己是站在局外人的立场上在看它。人们想知道，它到底有什么重大的影响？$E=mc^2$ 的情况就和社会名流一样，好像是由某种神秘的社会进程推动形成的。

然而，说到底名流指的是人，而 $E=mc^2$ 只是一个方程。它与其他方程一样，也是源于人们对现实世界进行描述的方式的不满。最初它的形式和现在人们看到的大相径庭。它修正了人们对于世界的看法，产生了意想不到的结果。

那么，这个方程到底是如何成为"名流"的呢？

牛顿与麦克斯韦之间的碰撞

方程的产生可以归结为几种不同的不满足感。有些不满足感来自于科学家的感觉，他们希望有一种新的方式能够对令人眼花缭乱的实验数据加以组织。另一些则源自对现有公式的不满，他们认为现有的方程太复杂了，需要简化，或者方程中的某些部分不协调。另外，理论预测与实验数据之间的误差所导致的不满，也会催生出新的方程来。

方程 $E=mc^2$ 的产生源于一种特殊且极为罕见的不满足感。19 世纪末 20 世纪初的许多物理学家都曾有过这种感觉。这种不满足感的产生最初是因为一个令人困惑的实验。这个实验指出了两种广为人知的伟大庄严的科学体系——牛顿体系和麦克斯韦体系之间的矛盾。更确切地讲，这个实验的结果指出了运动的相对性原理与光速恒定原理之间的冲突，而这两个原理分别是这两个科学体系的基础。



这一矛盾涉及一个称为不变性的原理。简单来讲，该原理说的是一种事物可以以两种截然不同的方式出现，但实质上仍是同一个事物。举例来说，两个人在同一房间的不同位置看同一把椅子，他们可以说椅子在电视的右边或左边。当考虑到两人位置的不同时，就能明显看出他们实际上描述的是同一把椅子，并且可以知道他们对椅子的描述会有怎样的不同，以及为何会产生这种不同。如果我们无法对这种差别作出解释，就会认为其中一人或两人都产生了幻觉，看到的是幻象。众所周知，对实际存在的事物来说，从不同的角度看会有所不同。因而，真相也必然会涉及本质与看到的表象之间的差别。换句话说，就是局部效应与整体性质的差异。观察物体时会看到物体的一个图景。这个图景会随着观察位置和光线等

条件的变化而变化。只要位置发生了改变，“局部”效应就会发生变化。然而无论位置如何改变，观察的始终是同一个物体。所以，不变性包含了对个体能够展示不同外表这一特性的理解。哲学家称此为虚实相关，物理学家则把它叫作变换下的不变性或协变性。简单来讲，协变性是客观性的定义的一部分。要说一个事物是真实世界的一部分，就得承认观察的角度不同，会导致所看到的图景也不同。但是，如果能用合适的变换加以描述，这些图景就可以殊途同归。

牛顿力学假定存在绝对空间和绝对时间。它们就像是竞技场或者舞台，所有事件都在“竞技场”和“舞台”上发生，不存在特殊的时间和空间。人们把这称为平移不变性，即物理定律不因时间和空间上的平移而改变。然而，牛顿原理指出了更深一层的不变性：根据“运动的相对性原理”，无论是在运动，还是静止状态下，都不存在特殊的运动。在匀速运动的系统中，物理定律都是相同的，与运动的方向和速度都无关。我们都有过类似的经验，比如在一列运行平稳的列车上，只要没有发生急刹车，人们的任何活动，如喝水、打牌、玩手球或跳舞都和列车停在站台上时一样。水安稳地装在杯子里而不溢出来，球在地板上同样的位置弹起，跳舞的人充满自信地在原地做出各种姿势。我们无法在火车上用实验验证列车行驶的速度，甚至无法验证列车是静止的还是运动的。而且，对于在另一辆静止列车上的人来讲，当考虑到两辆车之间速度的差异时，他们会看到和我们一样的物理定律。此处考虑两车速度的差异时所做的工作就是简单的相互加减。

科学家们把这辆（描述事件的立足点的）列车称为*参考系*，把匀速运动的参考系称为*惯性参考系*，把从一个参考系到另一个参考系变换事件的数学描述（位置坐标 x、y 和 z 以及时间 t）所用的方程称为*变换*，把联系不同惯性参考系中的对象性质的方程叫作*伽利略变换*。这种变换体现出的运动相对性原理早于牛顿力学，最先出现在伽利略力学中，伽利略晚期的假想实验就有从航船的桅杆上发射炮弹的。伽利略变换非常

简单，例如在行进的列车上，变化的量仅仅是火车相对于地面行驶的距离（将此称为 x 轴）。若相对火车的位置为 x'，火车行驶 t 时间后，与地面上观察者的距离为 x，那么有：$x'=x-vt$。另两个坐标 y 和 z 保持不变，所有事件在同一时刻 t 发生。

物理学家对真实性和客观性的定义依赖于伽利略变换。“真实”物体或事件是在不同惯性参考系中具有相同物理描述的物体或事件。只要运用适当的变换，将运动速度和方向上的差异考虑进去即可。真实性要求区分清楚事物的表象和描述间的差异，观察角度的不同会对所观察事物的客观性产生影响。在变换发展的过程中，科学家仅仅将重心放在了如何保持客观性上——无论从哪个惯性参考系看去，这种客观性都应保持不变。

因此，运动的相对性是牛顿力学的核心。但是，根据“光速不变原理”——麦克斯韦原理的核心概念，“光”为牛顿力学的简洁图景带来了新的元素。光与声音类似，无论声源的速度多大，声音总是以恒定的速度传播（在空气中声速约为每秒 1 100 英尺）。这是因为决定声音传播速度的是传播介质的性质（这里是空气分子），所以声波的传播速度无法超过某一上限。根据麦克斯韦方程组，光也总是以恒定的速度传播（约为每秒 186 000 英里），与光源的运动速度无关。为解决这个问题，物理学家假设光在一种被称为以太的介质中传播，以太的性质决定了光的传播速度。果真如此的话，这一概念将指明在绝对时间和绝对空间的“舞台”上，存在以太这一绝佳的惯性参考系。地球在以太中围绕着太阳公转，与此同时还可能“拖拽”些小型星体。地球相对于以太的运动速度可以通过检测不同方向光的速度之差来算出。由于以太改变了光的传播，因此这个速度差可以通过毕达哥拉斯定理计算出来。

试想如下场景，一艘快艇向着 400 码（1 码=91.44 厘米）外的对岸航行。在此过程中，水流将其冲向下游 300 码。当然，使船运动的载体是河水。因此，快艇沿着一个直角三角形斜边的方向运动，这个直角三

角形的两条直角边是指向对岸的 400 码的长度和指向下游的 300 码长度。快艇总的运动距离为（基于简单的毕达哥拉斯定理实例）500 码。为了能垂直驶向对岸，快艇应该以相同的角度朝着河的上游行驶。这最终将驶过更长的距离（直角三角形的斜边），但可以垂直抵达对岸。因为同样的原因，争议出来了：因为以太会改变光的运动，那么如果光的运动方向与以太运动方向垂直，光速就会发生变化。

1881 年和 1887 年，两位美国物理学家阿尔伯特·迈克尔逊和爱德华·莫雷进行了一项极为灵敏的实验，用来探测物理学家所称的“以太漂移”现象。实验装置包括两条“臂”，一条指向假设的以太运动方向，另一条与之垂直，光束能够沿着第二条臂来回传播。在水银槽上的石板上安装平面镜，并将其旋转 90 度，这样光线就能沿着与以太运动方向不同的方向传播。将两束光线靠近，就能产生干涉图案。据此迈克尔逊和莫雷就能够探测出光束速度上的微小差异。但是，该实验未能检测出这种差异。

物理学家感到很困惑，牛顿或麦克斯韦的方程组之中，肯定有不对的地方。

起初，他们认为问题出在麦克斯韦那里。毕竟他的理论形成较晚。麦克斯韦方程组出现不过几十年，而牛顿定律已存在两百余年，除某些存在微小的误差外，几乎成功解释了一切物理现象。但人们有理由倾向于认为此种结果是由实验误差或失误造成的。当时最聪明的物理学家试图通过修正麦克斯韦方程组，使它能与伽利略变换相符合。[2]但是，对麦克斯韦方程组进行改动是非常难的。它们包含由相互联系的原理构成的精巧网络。只要改动一个方程，由其他方程得出的结果就会很糟。

19 世纪末期，热衷于电动力学的物理学家们怀有一种深深的不满。他们认为必须对牛顿体系和麦克斯韦体系之间的悖论——光速不依赖于参考系的运动而保持恒定——作出解释，但是，没人能给出令人满意的答案。爱因斯坦曾经说过：“这个世界最让人不能理解之处就在于它是可

以理解的。”不言而喻，对于科学家来讲，最令人沮丧的事情莫过于无法解释这个世界了。

孤注一掷

不满逐渐发展成了绝望。1889 年，爱尔兰物理学家乔治・费兹杰拉德（George FitzGerald）写下了一段只有 5 句话而没有任何方程的文字，称能解释迈克尔逊–莫雷实验。这个能解决牛顿与麦克斯韦之间的冲突的“几乎唯一的假设”是：“随着物体相对于以太的运动，物体的尺寸会发生变化，变化量与光速的平方成正比。”[3]费兹杰拉德设想，在迈克尔逊和莫雷的装置中，指向以太运动方向的臂由于受到以太运动对其分子的作用而变短了。如果缩短的量恰到好处，那么用它测量出的光束沿着以太或背离以太运动方向的速度，与沿垂直臂方向运动光束的速度就相等了。然而，物体的尺寸在高速运动时会缩短的想法实在是很难让人接受。

另一个绝望的人是荷兰思想家亨德里克・洛伦兹（Hendrik Lorentz），1892 年时他在写给朋友劳德・雷利（Lord Rayleigh）的信中述说了迈克尔逊–莫雷实验带来的窘境。他写道：“对于这一矛盾，我完全摸不着头绪。”[4]那一年，他独立发表了与费兹杰拉德同样的想法，他说，“我能想到的该实验的唯一解释是”，以太的运动对刚性物体的长度产生了影响。洛伦兹得知费兹杰拉德的想法后，立即与他取得了联系。而费兹杰拉德也因遇到知音洛伦兹喜出望外，他坦言自己曾因这一想法而遭人“嘲讽”。[5]如果要解释长度收缩，就必须得到一组确切的变换，于是洛伦兹接着做了进一步的研究。在研究过程中，洛伦兹发现时间也会受到影响。当初，费兹杰拉德只是为了解释由迈克尔逊–莫雷实验得到的“沿着运动方向不同的光具有相同的速度”这一结论。洛伦兹则更加野心勃勃，想证明无论对于运动还是静止的观察者，光速均保持恒定。要实现这一目标，时钟就一定要变慢。之后，他提出了一组现在被称为洛伦兹变换的方程，对静止与运动系统中物体尺寸和时钟的之间的关系加以修正，以

解释迈克尔逊–莫雷实验中观察到的光速在以太中速度恒定不变的现象，进而解决牛顿和麦克斯韦理论间的矛盾。物体尺寸和时钟的修正因子均为 $\sqrt{1-v^2/c^2}$ 。[6]可以看出，如果没有相对运动（v 为 0），就不需要修正。在低速情况下，修正因子微乎其微，可以忽略不计。随着物体的运动速度增加，趋近光速时，修正因子就会显著增大，物体在运动方向上的尺缩效应越明显，时钟也会变得更慢。但是，当时大多数科学家都认为这种想法实在是匪夷所思，并没有放在心上。这也反映出当时普遍的事实是科学家们还在不惜一切代价地拯救以太。

正如后来一位科学家所说，这一切似乎是“大自然竭尽全力布置的阴谋”，为了“不让人们测量，甚至是探测以太中的运动”。[7]

恐慌不断加剧。一些著名的科学家开始提出充满幻想的理论。1898 年，法国著名数学家亨利 • 庞加莱（Henri Poincaré）为了支持“本地时间”的观念，提出了一种放弃绝对时间的理论，并通过该理论进一步解决了以太中的光速之谜。在 1904 年圣 • 路易斯举办的世界博览会上，庞加莱在公开演讲中几乎是带着戏谑的口吻说道：“或许应该提出一种全新的力学。这种力学我们只能够瞥见它的一瞬。在这种力学中，光速是永远无法达到的极限。”[8]

这样看来，被开尔文爵士称为遮盖了 19 世纪力学理论的“美和清晰”的“一号乌云”变得越来越难以理解了。庞加莱演讲的第二年，即 1905 年，为了解决这一难题出现了各种各样的理论，想要驱散这朵乌云，其中包括：高速状态下时间空间的收缩，绝对时间和空间的不存在，以及光速是速度的上限等。通过各种形式，这些理论都被证明是正确的。

爱因斯坦登场

当时只是专利局小职员的爱因斯坦为什么会关注这个问题？答案和其他人一样——不满。

多年后，就在爱因斯坦提出狭义相对论的概念已有五六个星期，但最终论文尚未完成之时，爱因斯坦在给朋友的信中写道：“(狭义相对论的）论证和基本架构经过了很多年的准备。”[9]

论证最早可以追溯到 1895 年底或 1896 年初，当时爱因斯坦 16 岁。这一思想是以他所谓的“天真的假想实验”形式提出的。(“天真”这一形容词常常被爱因斯坦使用，代表纯真、直接。）这个年轻人问自己：如果我以光速运动时，并去看旁边另一束并行的光线，会怎么样，又会看到什么呢？[10]牛顿说这可能会发生，麦克斯韦说不会。

我们能不能追上光波？这个问题虽然简单，但是也需要作出回答。不过答案无法用已有的物理学工具得出。这一问题激发了爱因斯坦的不满足感，困惑和好奇心驱使他去寻找问题的答案，提出论证，构建理论。

爱因斯坦为此苦苦思索了许多年。“必须承认，”他后来向一个朋友说，“狭义相对论的思想最初在我脑海中形成时，我被各种思想上的矛盾困扰着。年少时的我常常连着好几个星期陷入迷惑的状态之中，如同一个人在第一次遇到这个问题时，需要克服而又没能克服的麻木状态。”[11]1905 年的一天，爱因斯坦拜访了专利局的同事，也是自己的好友——麦克·贝索（Michele Besso）。他与贝索详细谈论了自己与这一问题“斗争”的历程，然后两人就道别了。然而，在阐述的过程中，爱因斯坦找到了问题的解决办法。第二天，他再次拜访贝索，见面就说：“十分感谢，我已经完全解决了这个问题。”[12]

结果就是 1905 年 6 月，爱因斯坦把论文“On the Electrodynamics of Moving Bodies”提交到了《物理学年鉴》。该论文是有史以来最著名、最重要的论文之一。虽然论文诞生的背后满是苦恼，但论文本身却遵循了一个简单有力的逻辑——“一种深刻的，几乎像孩子一样天真的新鲜方式”[13]，而且相对易于理解。

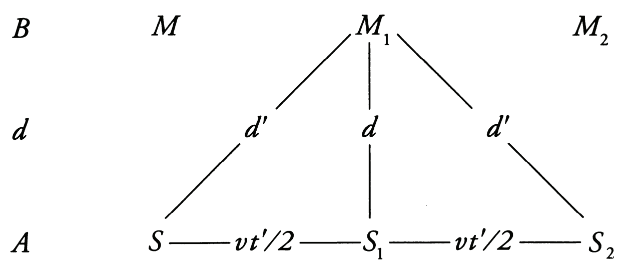
“众所周知，”爱因斯坦说，“现在一般人认为，麦克斯韦电动力学原理在应用到运动的物体上时，会得出与现象不符的结果（这些奇怪的结

果，更像产自理论的人为假设，而非真实世界中存在的）。”[14]他举例说，从“探测地球相对于‘光介质’的运动的失败尝试”出发，可以推出一个基本假设：“绝对静止”的概念是不存在的。他把这种假设称为“相对性原理”，并把该原理和另一个基本假定——“光波以恒定速度 v 进行传播，与发射光波的物体运动状态无关”——联系了起来。

因此，爱因斯坦的论文主要是围绕着解决两个重要原理的矛盾这一逻辑需要而设计的。这两个重要的原理就是相对性原理和光速不变原理。它们“好像又无法相容”，爱因斯坦说，但也仅仅是“好像”。在论文的其余部分，爱因斯坦只通过逻辑推理，就构造出一套“简单统一的动体电动力学”，而无需假定存在以太或其他绝对静止的参考系。怎样才能使两个不同的惯性参考系中的观察者得到同样的光速？爱因斯坦认为需要用到洛伦兹在物体运动方向上对空间和时间使用的修正因子。然而，洛伦兹（及费兹杰拉德）工作的基础是假定以太存在的，同时长度收缩也是的确存在的（以太对分子力的作用）。而爱因斯坦的工作仅仅以相对性原理和光速不变原理这两个假设为基础。也就是说，洛伦兹和费兹杰拉德必须假定以太是存在的才能得出结论，而爱因斯坦完全不用这一假设也能得出同样的结论。正如当时科学家们评论的：“掩盖阴谋之说是不存在的，因为压根儿就没什么可掩盖的。”或者就像费曼津津乐道的：“宇宙的阴谋就是自然定律。”

爱因斯坦在论文中用“β”代表修正因子，整个推导过程常以最简单的毕达哥拉斯定理的方式呈现。假设两个惯性参考系 A 和 B 以相对速度 v 运动。在惯性参考系 A 中，光源垂直于运动的方向发出一束光，照到距离为 d 的一面镜子上，并被反射回来。在惯性参考系 A 中观察，可以很容易地得出光传播的总距离为 $2d$。但是在惯性参考系 B 中观察，A 中所有的东西，包括光源、镜子，都在以速度 v 运动，因此光线也传播了更长的距离，为 $2d'$ 。半光程 d' 是直角三角形的斜边，两个直角边分别为 d 和 $vt'/2$，且有 $(d')^2=(d)^2+(vt'/2)^2$ 。然而，依照第二条原理，无

论在 A 还是在 B 中观察，光速都应该是 c，相同时间内应该传播相同的距离。从而，在 A 中 V（爱因斯坦用此符号表示光速）应该等于 $2d/t$，而在 B 中 V 等于 $2d'/t'$。怎样才能得出这样的结果？只有 A 中的距离和时间在 B 中看来同时变短才行。那么，要变短多少呢？按照 d 相对于 d' 变短的比例就行，也就是 d/d' 或者 t/t'，或者直接称为变换因子 β。如果 $V=2d/t$，则 $d=Vt/2$；如果 $V=2d'/t'$，则 $d'=Vt'/2$。将上式代入毕达哥拉斯定理方程，可以得到 β（即我们所求的比例因子 t/t'）为 $\sqrt{1-v^2/c^2}$。



这篇被称为“狭义相对论”（以区别于爱因斯坦于 1915 年发表的“广义相对论”）的重要论文发表于 1905 年 9 月 26 日。它颠覆了人们传统的时空观念。这是一篇有着如此重要意义的论文，特别是它的完成是作者在诸多领域热切工作的结果，这意味着作者在写作时一定有很多难以预料的结果。下面这个结果就来得极其迅速。1905 年秋季的一天，爱因斯坦在给好友康拉德·哈比希特（Conrad Habicht）的信中写道：

> 研究电动力学的一个结论掠过了我的脑海。这就是与麦克斯韦基本方程相联系的相对性原理，要求物体的质量为它所包含的能量的一个直接度量，光也有质量。镭是唯一能观察到的质量减少的例子。这种想法既有趣又迷人，可我知道，全能的上帝对这一切或许嗤之以鼻，或许正是他在牵着我的鼻子转呢。[15]

这种像美诺的奴隶感受到的那种“被牵着鼻子走”的感觉使爱因斯坦认识到有些看上去是对的，但也还需要进一步研究。

相对论论文发表后的第二天，爱因斯坦又以“物质的惯性依赖于其能量吗？”（Does the Inertia of a Body Depend Upon its Energy Content?）为题向《物理学年鉴》寄出了一篇三页纸的论文，专门探讨上述问题，这篇论文也于同年发表。科学史学家约翰 • 里格登（John Rigden）指出，这篇论文并没有开辟出新的领域，只是提出了一个之前的论文中逻辑上还不是很清楚的结论。这一结论本来就可以很自然地成为前一论文的最后一部分。里格登说，果真如此的话，“它将会是一个了不起的结论。”[16]

在上文的开场白，爱因斯坦以一种谦逊的口吻说：“从本刊新近登出的本人在电动力学方面的研究结果，可以得出有趣的结论。”爱因斯坦通过下面的例子推导出了结论。假设在参考系 A 中，处于静止状态的一个物体（比如一个原子）的质量为 m。m 向着相反的方向发射出两束光线（即释放能量）。我们假设总的能量损失为 L（和以前的论文一样，爱因斯坦用 L 表示能量，V 表示光速，现在这种记法已经不常用了），那么每束光所携带的能量均为 $L/2$。A 中的观察者认为物体的动量没有发生变化，原子仍然处于静止状态。原子将激发态的一些能量释放出来，质量与之前相比没有变化。但对于参考系 B 中的观察者来说，A 是运动的，他看到的情况会有些不同。向前发射的光束比向后发射的光束动量大。这意味着原子的动能有净的减小。只有原子的速度或质量减小时，动能才会减小。而原子速度不变；并且在静止参考系中，没有尺缩效应。唯一其他的可能性就是在运动的参考系中，原子的质量减小了。在原子所在的静止参考系中来看，原子的质量没有增加，“惯性”没有变化。而在假想的运动参考系中来看，原子的“惯性”变化了。那么变化量是多少呢？爱因斯坦用前面论文中的方法算出了转换因子，仍为 β。

爱因斯坦仍然使用现在已不常用的记号 L 和 V，继续写道：

> 如果物体通过辐射释放了能量 L，它的质量减小 L/V^2。此处很明显，从物体中发射出的能量是转化成了辐射能量还是其他某种形式的能量并不重要。我们可以得出一个更一般的结论：物体的质量是其能量的度量。[17]

这就是著名方程 $E = mc^2$ 出现之前的最初形式。它虽然没有明确采用方程的形式，也没用采用人们熟悉的符号，但已经清晰地体现了具有划时代意义的质量–能量原理。这一概念改变了宇宙结构等一些最基本的概念。它将两个长久以来被认为是完全不同的概念——能量和质量，联系在了一起：能量守恒原理是 19 世纪物理学上最耀眼的成就，而质量守恒定律则是 18 世纪科学上的耀眼成就。[18]两者可以相互转化。

这一理论变革了对客观性的理解。在牛顿时代，能量和质量在不同的惯性参考系中保持不变；在爱因斯坦时代，在低速情况下质量和能量近似保持不变，而在速度接近光速时，二者将发生变化。客观地说，从另一个速度足够快的惯性系观察时，长度和时间在以这样的比例发生变化。[19]

接下来的几年里，爱因斯坦多次提到这一结论，不过仍然是以描述的方式或使用自己的符号，而不是现在人们熟悉的形式。如在 1906 年一篇论文的脚注里，爱因斯坦写道：“质量守恒是能量守恒的一种特殊形式。”[20]1907 年初，爱因斯坦在另一篇发表在《物理学年鉴》上的论文里用 ε 代表能量，希腊字母 μ 代表质量，V 代表光速，得到下面的方程：

$$\varepsilon = \mu V^2 \frac{1}{\sqrt{1-(v/V)^2}}$$

这一著名方程（能量等于质量乘以光速的平方）中含有修正因子 β，考虑了物体运动时的效应。以电子为例，在静止状态下，所有电子的质量都相同。这个质量是电子的固有属性，在电子产生时就有了。在电子自身的参考系中，不论何时对电子进行称量，电子的质量都保持不变。现

在我们假设从一个运动的参考系中观察该电子。如果 $E = mc^2$ ，且 c 为常数，那么随着能量的增加，m 和 E 应该具有相同的变化趋势。电子的惯性质量（电子在自身静止参考系中的质量）保持不变。但在实验室中测定的运动电子的质量会有所不同。修正因子 β 告诉我们应该乘以多大的系数，才能得到电子的惯性质量。如果把修正因子项去掉，就可以得到爱因斯坦在脚注中提到的“简化公式 $\mu V^2 = \varepsilon_0$”。[21]

同一年的晚些时间，爱因斯坦将光速的符号由 V 变成了 c。相对论理论所包含的结论具有“极其重要的理论意义”，他说：“在相对论中，物理系统的惯性质量和能量是一回事。提到惯性，惯性质量为 μ 的物体等价于能量 μV^2。”[22]在接下来的几年，爱因斯坦得出了更为完整的质能原理和推论。在 1912 年的一篇有关相对论的手稿中，爱因斯坦在开始讨论质能原理这一论题时，用符号 m 代替了原来的 μ，用 L（和最初的写法一样）代替 ε，然后又划去，改为 E。从此，他一直坚持使用 E 和 c，于是便有了下面这个包含了修正项的广为人知的方程，其中 q（有时写作 v）表示速度：[23]

$$E = \frac{mc^2}{\sqrt{1 - q^2 / c^2}}$$

进入原子核

每个重大科学发现都不可避免地会引起一些问题：为何以前没有发现这一现象或原理？答案往往比较复杂，影响因素很多。一个因素是科学家们虽然遇到过这个问题，可是把它忽略或曲解了，或者是对这一问题还没有完整的描述。质能转换原理就是这样。另一个因素是已有的科学知识阻碍了人们对新现象和新原理的研究。质能转换原理也包含了这一因素。以前人们都把质量和能量视为两类性质完全不同、遵循不同定律的量。最后一个因素是大多数的现象或原理的表现形式都不是能够容

易地被科学家探索和研究的。这些现象或原理的效应往往是很微小的。这一点在质能转换原理中也有体现，人们在日常生活中无法察觉到。爱因斯坦写道：“这就好比有一个人，虽然腰缠万贯，却向来一毛不拔。你怎么知道他有多么富有呢？”[24]

那么这个有钱人到底会不会花钱？花在哪儿？在“能量”一文中，爱因斯坦采用一种更为慎重的方式表达了自己的观点，而不像在给哈比希特的信中那么热情洋溢。爱因斯坦写道：“或许可以用像镭之类的以辐射的形式释放能量的物质去验证该理论。”但是，爱因斯坦不久后在另一篇论文中提到，这种效应“微乎其微，无法测到”。他引用普朗克的计算结果，指出镭的质量损失所需的时间超出了“实验可接受的范围”[25]。他接着写道：“与镭不同，如果原子最初的能量能有很大一部分转化为各种形式的辐射能，就可能会探测到放射性过程。”

1911 年原子核的发现对验证质能原理并没有起多少作用，在之后的 20 多年里也是如此。然而 1932 年的两个关键性进展，却使质能原理在理解宇宙时变得非常有用，甚至成为从宏观到微观解释宇宙（小到原子结构，大到恒星爆炸）所不可缺少的工具。第一个进展是英国物理学家詹姆斯·查德威克发现了中子。如今，物理学家已经对原子核的基本结构有了较好的认识：原子核包含质子和中子。那么是什么力量使质子和中子结合在一起的？第二个关键性的发现为此提供了线索。这一发现是由另两个英国物理学家科克罗夫特（John Cockroft）和沃尔顿（Ernest Walton）于 1932 年作出的。他们使用了一种叫作粒子加速器的新设备对质子进行加速，轰击锂核，引发核反应：一个锂核加上一个质子变成两个氦核。科克罗夫特和沃尔顿测量了始态（锂核和质子）的质量和能量，又测量了末态（氦核）的质量和能量，发现总的质量损失和能量增长的关系，在实验误差允许的范围内与爱因斯坦的质能方程完全吻合。惯性质量的减小量等于动能的增加量除以光速的平方。这是爱因斯坦的质能方程第一次被证实，并且很快在原子物理领域开始发挥无可替代的作用。

同种粒子在原子核内外的质量之差称为“敛集系数”。所有这类粒子分别位于原子核内外时的总质量差称为结合能。与此同时，物理学家也了解到星体所发出的光来自原子内部的质能转化过程。20 世纪 30 年代，敛集系数和结合能的概念使爱因斯坦方程成为科学界中的一个实用工具，从原子物理到天体物理皆是如此。

物理学家意识到，即便只有很小的质量发生转化，都能产生极大的能量。所释放出的能量将超过迄今为止人类所知道的任何一种物理过程。然而，单个原子核即便是释放出所有的结合能，所产生的能量也是微乎其微，没有实用价值。因此，核能在此后的十年都被认为是遥远的甚至是荒谬的想法，无非是痴人说梦。直到 30 年代末期，几乎所有的科学家都还认为释放和控制核能是牵强的甚至是疯狂的。1921 年，有个年轻人想通过 $E = mc^2$ 预示的强大威力制造出一种新式武器。为此这位年轻人来请教爱因斯坦。爱因斯坦回答说:“一眼就能看出这想法愚蠢至极。”[26] 1933 年，在一次采访中，物理学家卢瑟福（Ernest Rutherford）称这种想法是“痴心妄想”。1936 年，丹麦物理学家尼尔斯·玻尔（Niels Bohr）在讨论粒子与原子核碰撞会释放出核能的这一现象时，说这并不会“把老生常谈的核能实用化问题向前推进”。玻尔还说道：“的确，对核反应的了解越多，似乎这一目标离我们越远。”[27]

然而，当时已有一系列的事件初见端倪，并将改变世界对质能转换的看法。这些现已广为人知的科学和政治事件接二连三，人员阵容遍布全球，并且带有鲜明的戏剧性。即便在半个多世纪后的今天看来，即便只是看一下概要，也能令人激动不已。

1932 年查德威克发现中子后，科学家们立即意识到这种粒子是研究原子核的极佳工具。20 世纪 30 年代中期，正当法西斯在欧洲叫嚣之时，意大利物理学家恩里克·费米（Enrico Fermi）用氦核从头到尾依次轰击元素周期表中的元素，分别得到了各元素的较重的、具有放射性的同位素。当他的工作进行到当时已知的最重元素——铀时，奇怪的结果出现

了。费米认为自己得到了全新的“超铀”元素。

德国科学家奥托·哈恩（Otto Hahn）和弗里茨·斯特拉斯曼（Fritz Strassman）发现费米错了，用氦核轰击铀得到的实际上是一种人们熟悉的更轻的元素。1938 年 11 月，他们将这一结果寄给了以前一起工作过的同事——里瑟·迈特纳（Lise Meitner），迈特纳当时为躲避纳粹逃到了瑞典。迈特纳和自己的侄子，也是物理学家的奥托·弗里希（Otto Frisch）认为轰击的结果实际上是分裂了原子核。在咨询了一个生物学家之后，弗里希将此现象命名为“裂变”。随后，弗里希和迈特纳写成了一篇在核裂变方面具有里程碑意义的论文，并寄给了《自然》杂志。该杂志在 1939 年 2 月刊登了这篇论文。但在论文登出前，弗里希已经将论文的内容告诉了玻尔。当时玻尔正要坐船赶往美国。一到美国，玻尔和同事就在一月中旬普林斯顿举行的物理系月度聚会上将核裂变的消息告诉了美国的物理学家们。一周后，哥伦比亚大学的物理学家们开展了美国本土上的第一次核裂变实验，消息迅速传遍全国。科学家们最先是从报纸上，而不是从物理期刊上获知这一消息的。大部分科学家都知道到在铀裂变的过程中，会释放出一个氦核。这个氦核会引起另一个铀的裂变，由此产生的链式反应会引起大量的铀在瞬间发生裂变，释放出巨大的能量——这也就意味着人们能够据此造成一种新式的、威力惊人的炸弹。而此时的欧洲，大战一触即发。

1939 年 3 月，费米（此时为了躲避法西斯的迫害，已经从意大利逃到了美国，先是在哥伦比亚大学，后转至芝加哥大学）和其他物理学家开始正式向美国政府报告核裂变在军事上可能的应用。7 月间，两位科学家拜访了正在长岛的皮科尼克家中避暑的爱因斯坦，向他寻求帮助。在得知链式反应的可能性后，爱因斯坦大呼：“我从来没想到过这个！”两周后，他签署了一封紧急信件，递交给了总统罗斯福，信上说道：“最近的工作使我想到，铀元素在不久的将来可能会成为一种新型重要能源。”在接下来的 4 个月里，通信一直在继续，不过信中已经开始讨论利

用铀进行链式反应的可能性。“这一新的现象也可以用于制造炸弹。虽然不确定性很大，不过可以预见由此可能会造出威力极大的新型炸弹来。”

1939 年 9 月，纳粹德国入侵波兰。当年 10 月，爱因斯坦上书美国总统罗斯福，建议研制原子弹。1940 年 2 月，联邦政府拨款 6000 美元，成立了曼哈顿计划（Manhattan Project），用于研究核反应现象。包括德国、前苏联、日本和英国在内的多个国家相继开始了原子弹的研究，但只有美国的研究工作进展迅速。

1942 年 12 月 2 日，在计划正式开展不到一年的时间，世界上出现了第一个可控链式反应。地点位于芝加哥大学足球场西侧壁球场内的冶金实验室，这一成果证实了曼哈顿计划的可能性。[外界的消息是：“意大利领航员（费米）在新大陆着陆了。”] 为此，罗斯福总统为计划又追加了 40 万美元拨款，在田纳西的橡树岭建立大型同位素分离工厂，并在华盛顿的汉福德建立钚生产工厂。J. 罗伯特 · 奥本海默（J. Robert Oppenheimer）担任该计划首席科学家。他选定新墨西哥的洛斯阿拉莫斯（Los Alamos）作为原子弹的实际研发基地。该基地坐落在洛斯阿拉莫斯山顶，位置偏僻，比较安全。从 1943 年 3 月起，科学家们陆续进驻此处。

1945 年 7 月 16 日，在美国新墨西哥州的洛斯阿拉莫斯进行了一次试爆，标志着曼哈顿计划达到了顶峰。科学家们惯于在实验室观察新的现象，但这次核试验却大不相同。在沙漠中的寒冷清晨，洛斯阿拉莫斯的科学家们蹲在地上，手持焊工的护目镜遮住双眼，静静等待。突然，一团火球腾空而起，比太阳还要刺眼。远在 20 英里之外的科学家们也能感受到火球放出的热量。随后又慢慢升起了数万英尺高的白云。此种场面令科学家们不禁担心起来，他们释放出了自己无法控制的巨大能量，这让奥本海默想到（他后来说的）圣经中的天启之路。阿布拉罕 · 佩斯（Abraham Pais）写道，这是“人类历史上最为壮观的场景”。[28]

3 周后的 1945 年 8 月 6 日，世界上第一颗原子弹将日本的广岛市夷为平地。第二天，世界各地的报纸头条都竞相报道这一极可怕的新式炸

弹。《纽约时报》评论说：“这是爱因斯坦教授的相对论第一次走出实验室，得到实际应用。”[29]8 月 9 日，就在广岛原子弹爆炸后的第三天，另一颗原子弹又摧毁了长崎市。

方程 $E = mc^2$ 本身并没有在促成曼哈顿计划的过程中发挥直接的作用，只是在理解核裂变现象时提供了物理依据。原子弹通过裂变实现了质量到能量的转化。它只是方程 $E = mc^2$ 的一个应用——一个地球生命来去轮回的少有应用而已，而非方程本身导致了原子弹的出现。不过，曼哈顿计划的一位成员、普林斯顿大学的物理学家亨利 · D. 史密斯（Henry D. Smyth）所撰写的一篇报告《原子能在军事上的应用》（*Atomic Energy for Military Purpose*）却还是把质能方程直接和原子弹联系到了一起。

8 月 11 日，在长崎原子弹爆炸后的第三天，史密斯的报告被公之于众。“破坏力远超出人类想象的武器已被制造了出来，”史密斯写道，“它不是由充满邪念的畸形天才创造出来的，而是由千百个为了维护祖国和平而辛勤劳作的人们创造出来的。”本来这篇报道的预期读者是“工程师和科学工作者们”，这些人可以“向民众介绍原子弹的潜能”。然而该书却出人意料地大受追捧，头 5 个月销量就突破了 10 万册。[30]

报告一开始，史密斯就采用 $E = mc^2$ 作为解释原子弹的基础。报告说狭义相对论的最早意义就是在质量和能量之间的等价性。

> 对大多数真正的物理学家和工程师来讲，这只是一个没有实际应用价值的数学上的幻想。连爱因斯坦本人也很难想见质能原理现在的应用。不过早在 1905 年，他就明确指出质量和能量是等价的，并提出或许可以通过放射性物质的研究找到对这一关系的证明方法。他总结出能量 E 与质量 m 的等价关系方程：
>
> $$E = mc^2$$
>
> 其中 c 为光速。

通过史密斯的报告，公众了解了曼哈顿计划。这份独一无二的报告使 $E=mc^2$ 成为了原子能和核武器的象征。

重要地位

民族学者指出，两种文明相互影响时，并不是浑然一体地相互融合，而是通过一种文明中的某些成员观察所用的“同源物”（congener），来理解另一种文明，并作出应答。这些同源物包括人造物品、宗教仪式、风俗和艺术等，恐惧、魅力和异国情调常常也会起到作用。同源物就像一个放大镜，人们通过它去观察其他文明，获得一个大致的了解。所以，同源物不仅是其他文明的符号或标志，也能在人们初次接触时指导和规范对于该文明的好奇心与迷恋。

从这种意义上说，方程 $E=mc^2$ 在公众对核能信息的渴求和让原子能成为现实的科技发展之间扮演了同源物的角色。在这一过程中，$E=mc^2$ 进一步又成为物理学、科学甚至是知识的象征，已经到了具有传奇色彩的地步。

法国智者罗兰·巴特（Roland Barthes）写过一篇关于爱因斯坦的文章，提到爱因斯坦的照片常常是他站在黑板前的形象，背后的黑板上写满了各种匪夷所思的符号和方程，而在漫画中，爱因斯坦则手拿粉笔，也站在黑板前，但黑板上只有方程 $E=mc^2$，就如凭空想出的一般。巴特注意到，这个方程的意义在于它向人们传达了这样一种信息：“知识被简化成方程……科学用几个字母就表达出来了。”它已经变成了斯诺底圣像：“自然界的统一、世界最基本的简化的理想状态、文字之中未释放出的能量、秘密与表述间的古老争斗以及全部知识瞬间就可以发现的想法，就像一把锁，在试了上千次之后一下子打开了。”巴斯的文章有助于解释方程 $E=mc^2$ 从科学工具到同源物的转变。

爱因斯坦本人也开始采用这个方程的简化形式。这一简化形式现已人所共知。1946 年 4 月，新近流行的杂志《科学画刊》（*Science Illustrated*）

在创刊号上有一篇题为“$E = mc^2$”的文章，作者是爱因斯坦。他写道，“用公式 $E = mc^2$ 表示能量与质量的等价关系（尽管不是很准确）是惯常做法。”[31]

1946 年 7 月 1 日，广岛和长崎原子弹爆炸后不到一年，也是 $E = mc^2$ 以最初形式诞生即将 41 年之际，它登上了《时代》杂志的封面。该期杂志发行时恰逢南太平洋的一次核试验。杂志的封面上是已经 66 岁的、白发苍苍的、被人们称为“腼腆、几乎如圣徒般、颇具孩子气的小老头”的爱因斯坦。在头像旁边，一团蘑菇云从一艘艘战舰上方升起。底部的红焰为橘色和紫色的柱状物所取代。柱状物的顶部是灰色的蘑菇云。云团上写着著名方程 $E = mc^2$。现在它已然成为方程中的名流。

 茶 歇

疯狂的想法

> 为了能理解这一点，我们需要些疯狂的想法。谁有疯狂的想法？
>
> ——尼尔斯 • 玻尔

科学家和科普作家杰里米 • 伯恩斯坦（Jeremy Bernstein）说他偶尔会幻想如下场景：

> 1905 年，当时我是波恩大学的物理学教授。电话铃响了，对方自称是一个我从没听说过的瑞士国家专利局的专利审查员。他说听说我在教电磁理论，自己有一些我或许会感兴趣的想法。“什么想法？”我略带点傲慢的口气问道。他开始谈论一

> 些关于空间和时间的疯狂想法：尺子在运动时会缩短，赤道上的钟要比位于北极的同样一只钟走得慢，电子的质量随着运动速度的增加而增加，两件事情是否是同时的取决于观察者所在的参考系等，不一而足。我该作何反应？[32]

伯恩斯坦的这个“假想实验”反映了民间科学家的一大特点：他们向别人写信宣传自己的理论时总是很疯狂。以前，这些来信都装在棕色的信封里，信上全是歪歪扭扭的手写体；现在则是附有俗丽网页链接的电子邮件。信的主题通常是天体物理学、宇宙学、统一理论和对西方科学的颠覆。爱因斯坦常会受到牵连，不是被当作主流科学的代表（也就是作者的敌人），就是被当作受到误解的主流科学外的孤独行者（又成了作者的前辈）。一般情况下涉及的是如下一个或几个主题：万有引力，电磁学和行星轨道；最原始的很容易发现的版本也常会提到超自然现象、星座、中药、股票市场、棒球得分和摇滚歌词等。许多狂想的信件常常满是*斜体字*、**加粗字**和**大写字母**，很容易让人联想到右派的新闻通信和软件授权协议。有些作者对政府或科学界镇压他们思想的阴谋加以警告；而另一些作者则以慷慨示人。[33]

很少有人在收到疯狂想法信件后予以回复。人们认为这不仅可能会帮倒忙，而且甚至可能是危险的，愈发使作者感觉被误解，并收到更多的“紧急求助”。收信人一般只是很快地浏览一下信件，然后就扔进了“狂想信件”的抽屉中。但我几乎没遇到把信扔掉的人。

为什么不扔掉呢？

我的一位同事把他的装有“狂想信件”的抽屉与隔壁的艺术展作了一个比较。如果你足够耐心周密，还是能发现一些有价值的东西的。但是这一寻找的过程所需的时间太长，几乎是不可能完成的。其他人给出了一些心理学上的解释：人们敬仰甚至是嫉妒“狂想信件”的作者有着如此旺盛的精力和热情。我们会感到有一种秘密的共鸣——当我们拥有

一个被误解的事实的时候，不都这么觉得吗？说得再阴暗一点，我们很害怕读这些信，这就好像是看着精神火车出轨一般。但其他同事会把这些信保留下来，他们告诉我说因为“你永远不知道会发生什么……”

这些疯狂的信件有着有趣的哲学、心理学和社会背景。对这些疯狂信件和请求进行分类比看上去要困难许多。伟大的科学家同样也会变得不可思议：读者可以回想一下爱因斯坦对统一场论的痴迷和鲍林对维生素 C 的吹捧。然而，科学家们难道不喜欢“疯狂的想法”，不喜欢依赖于它们吗？我们不是都听过那个著名的故事：沃尔夫冈·泡利（Wolfgang Pauli）指责玻尔坚持疯狂的理论，而玻尔的回答是，他的理论的问题就出在“还不够疯狂”上，需要一个“更疯狂的想法”。

最后，说几个小故事，说明人们不应对于自己辨别疯狂信件的能力过于自信。

第一个故事是 25 岁的拉马努金。拉马努金是印度人，他于 1913 年写了一封信给几位英国著名的数学家。与其他几位数学家一样，哈代一开始把这封信当成了一封狂想的信，扔到了一边。之后他再读的时候，发现了这封信乃是出自天才之笔，很快就请拉马努金到英国去。在那里，拉马努金被公认为是人类历史上最杰出的数学家之一。决定了谁会受到眷顾的常常不是人们满意与否，而是人的偏见。

第二个故事是尼古拉斯·克里斯多费罗斯（Nicholas Christofilos）。克里斯多费罗斯在希腊的一家电梯安装公司做电气工程师，他的兴趣是粒子加速器。1949 年，克里斯多费罗斯写了一封手稿，提出了一个新的机制，寄给了伯克利的物理学家。他收到了回信，信中指出了手稿中的一些缺陷。克里斯多费罗斯对手稿加以修改，申请了美国专利，把修改稿又寄给了伯克利的物理学家。但这次信件却被置之不理。1952 年，在读到美国物理学家“发现”了一种原理与他的方法完全相同的新型加速法之后，克里斯多费罗斯联系了一家法律机构，使自己的优先权得到了认可。我曾问一位看过克里斯多费罗斯原稿的科学家为什么会忽视他的

手稿。这位科学家说："第一是违反了麦克斯韦方程。"然后只是耸了耸肩，没再多说。他的肢体语言告诉我们这就好比是在说灵异现象一样，所以他也自然无需为忽视了克里斯多费罗斯以后的其他手稿而道歉。由此可见，糟糕的物理学并不一定就是狂想。

伯恩斯坦坚持认为，如果自己看到了通信者新近发表论文（动体的电动力学）的复印件，就会指出这并不是一篇问题论文。他还援引了两个特点。第一个是"联系性"，亦即当动体的速度与光速相比较低时，从该理论出发得出的结论与牛顿力学相同。而问题理论常常"始自半空，终至半空"，令人摸不着头脑，不能与现有科学知识体系联系起来。第二个特点就是已有可测试的预测结果。

我还要再加上两条特点。第一条就是作者处理方程的方式。疯狂信件几乎总是要么没有方程，要么把一个小数字当成崇拜的对象。的确，这些信中的方程通常会体现出巴特所发现的把知识简化为公式的理性的幻想。方程常常是单独出现的，就好像它们是独立存在的一些事实。人们把方程当作原理一样对待，把整个哲学纳入小小的方程之中。这篇论文里的方程就像是一个人拿着乐器却从未用它来演奏过。相反，在真正的科学论文中，方程几乎从来都不是单独出现的，而是作为冗长的逻辑推论中，依次出现的元素中的一员出现的。它们是知识大厦的一部分。有了知识大厦，方程的存在才有意义。而印在纸上的只是知识大厦的一小部分。也就是说，处理方程时不能把它们当成是完全独立的。

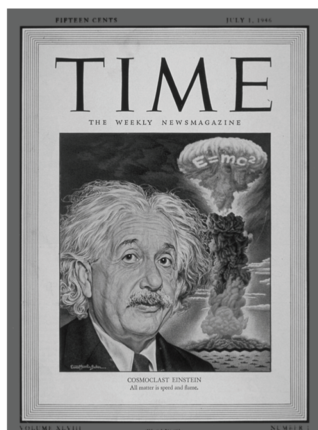
但我认为疯狂信件最重要的特点是缺少专注的态度和玩兴。爱因斯坦在向哈比希特提到上帝在牵着自己鼻子走，而他却情愿走下去的时候，在自己的担心中也表露出了这种态度。

作为例证，再举一例。这个故事发生在 1946 年 9 月在纽约城召开的战后美国物理学会第一次年会上。会上，一位年长的维也纳物理学家菲利克斯·厄仑霍夫特（Felix Ehrenhaft）打断了年轻的荷兰理论家阿伯拉罕·派斯（Abraham Pais）的演讲。派斯正努力地解释新近发现的一种新

粒子的令人困惑的奇怪行为。厄仑霍夫特自 1910 年就一直声称有“亚电子”存在（亚电子所带电荷比电子要小）的证据。他为推动自己的观点所作的努力长久以来已经耗尽了物理学家们的耐心。现在厄仑霍夫特即将步入 70 岁高龄，可他仍在寻求听众。于是他走到台上，要求说几句。

告诉我这个故事的是一位叫做赫伯特·戈尔茨坦（Herbert Goldstein）的年轻物理学家。他当时与导师也是以前在 MIT 辐射实验室的同事阿诺德·西格尔特（Arnold Siegert）坐在一起，他问西格尔特：“派斯的理论比厄仑霍夫特的要疯狂多了。可是为什么我们把派斯称为物理学家，却把厄仑霍夫特视为顽固不化之徒呢？”

西格尔特想了一会，坚定地说：“因为厄仑霍夫特过于相信他的理论。”



西格尔特的意思其实是厄仑霍夫特顽固不化的程度已经影响到了科学家所需的玩的态度。这种态度使科学家在带着不满足感前行时能够去冒险和回应。（尼采曾说：相比谎言，顽固是真理更大的敌人。）人们把某些思想视为疯狂思想，并不单是因为人们心存偏见，也不一定是个人观点，而是由于认识到了作者顽固不化的破坏效应。顽固不仅会扫除人们的不满足感，也会扫除人们的玩兴。只有把不满足感和玩兴结合起来，才能在科学上产生强大的推动力。

注　释

[1] Luce Irigaray, *Parler n’est jamais neutre*（巴黎：Editions de Minuit，1987 年），第 110 页。

[2] 有关麦克斯韦某些修饰符的简要讨论，参见 Alfred Bork，“Physics Just Before Einstein”，《科学》152（1966），第 597～603 页。
[3] G. FitzGerald，“The Ether and the Earth’s Atmosphere”，《科学》13（1889），第 390 页。
[4] Lorentz 于 1892 年 8 月 18 日致 Rayleigh 的信，引自 John S. Rigden，《爱因斯坦 1905：伟大的标准》（*Einstein 1905: The Standard of Greatness*，剑桥，马萨诸塞州：哈佛大学出版社，2005 年），第 82 页。
[5] G. FitzGerald 于 1894 年 11 月 14 日致 H. Lorentz 的信，引自 Abraham Pais《上帝是微妙的：爱因斯坦的科学与生平》（*“Subtle is the Lord”: The Science and the Life of Albert Einstein*，纽约：牛津大学出版社，1982 年），第 124 页。
[6] 我认识的一位物理学家是这样记起来静止参考系中的时间膨胀的：“宇宙射线可以到达地球。”也就是说，在静止参考系中，宇宙射线的寿命通常太短，无法传播很长的距离。但在地球看来，它们以接近于光速的速度在传播的，所以时间膨胀足以保证它们到达地面。
[7] 爱丁顿，“Gravitation and the Principle of Relativity”，《自然》，第 101 卷，1918 年，第 15~17 页（所引部分在 16 页）。
[8] 派斯，《上帝是微妙的》（*Subtle is the Lord*），第 128 页。
[9] Carl Seeling 于 1952 年 3 月 11 日致爱因斯坦的信，引自 Ronald W. Clark，《爱因斯坦：生活与时代》（*Einstein: The Life and Times*，纽约：世界出版公司，1971 年），第 84 页。
[10] P. A. Schilpp 编，《爱因斯坦，哲学家与科学家》（*Albert Einstein: Philosopher-Scientist*，伦敦：剑桥大学出版社，1970 年），第 53 页。
[11] 引自 Clark，《爱因斯坦》（Einstein），第 84 页。
[12] 引自派斯，《上帝是微妙的》（*Subtle is the Lord*），第 139 页。
[13] Emilio Segre，《从 X 射线到夸克》（*From X-rays to Quarks*，纽约：多佛出版社，1980 年），第 84 页。
[14] 爱因斯坦，“On the Electrodynamics of Moving Bodies”，《物理学年鉴 17》（*Annalen der Physik 17*，1905），第 891~921 页。该文收于爱因斯坦的《爱因斯坦论文集，第 2 卷，瑞士岁月：著作，1900~1909 年》（*The Selected Papers of Albert Einstein, Vol. 2, The Swiss Years: Writings, 1900~1909*），A. Beck 译，普林斯顿：普林斯顿大学出版社，1989 年，文档 23，第 140～171 页。
[15] 爱因斯坦于 1905 年 6 月 30 日致 Conrad Habicht 的信。收在《爱因斯坦论文集》中，第 5 卷，第 20~21 页。

[16] Rigden，《爱因斯坦 1905》，第 112 页。

[17] 爱因斯坦，《爱因斯坦论文集》，第 2 卷，文 24，第 174 页。

[18] 阿伯拉罕·派斯这样描述爱因斯坦的成就：“爱因斯坦在物理学上的伟大创新有两点。第一，记录空间间隔和持续时间的测定结果需要比当时已有技术更加精细的技术条件；第二，经典物理学只有在 $v/c \ll 1$ 的情况下才适用。化学上的伟大创新在于：拉瓦锡的质量守恒定律和道尔顿的重量成比例定律虽是近似的，却很实用，在传统化学中不需引入任何变化。于是，相对论把牛顿力学和经典化学变成了近似的科学。这两者在该过程中非但没有消失，反而得到了更好的定义。”Pais，《上帝是微妙的》（*Subtle is the Lord*），第 163 页。

[19] 这又是客观性或者协变性要求在客观性的普通概念和科学概念之间加入一个新概念的例子。

[20] 爱因斯坦，“引力中心运动和能量惯性的守恒原理”（The Principle of Conservation of Motion of the Center of Gravity and the Inertia of Energy”，《物理学年鉴 20》（*Annalen der Physik 20*，1906），第 627~633 页。收在《爱因斯坦论文集》，第 2 卷，第 200~206 页。

[21] 爱因斯坦，“论相对论原理要求的能量惯性”（On the Inertia of Energy Required by the Relativity Principle”，《物理学年鉴 23》（*Annalen der Physik 23*，1907），第 371~384 页。收在《爱因斯坦论文集》，第 2 卷，第 249 页。

[22] 爱因斯坦，“论相对论原理和从相对论原理得出的结论”（On the Relativity Principle and the Conclusions Drawn from It”，《爱因斯坦论文集》，第 2 卷，第 286~287 页。

[23] 爱因斯坦，《爱因斯坦 1912 年论狭义相对论的手稿》（*Einstein's 1912 Manuscript on the Special Theory of Relativity*，纽约：Braziller，1996 年），第 102~103 页，109 页。

[24] “爱因斯坦，$E=mc^2$：现时代最紧要的问题”（A. Einstein，$E=mc^2$: The Most Urgent Problem of Our Time），《科学画报》（Science Illustrated），1946 年 4 月，第 16~17 页。

[25] 普朗克，引自爱因斯坦，《爱因斯坦论文集》，第 2 卷，第 287 页。

[26] Clark，爱因斯坦，第 101 页。

[27] 玻尔，《自然》，1936 年 2 月 29 日。

[28] 阿伯拉罕·派斯，《奥本海默的一生》（*J. Robert Oppenheimer: A life*），Robert P. Crease 提供补充材料，纽约：牛津大学出版社，2006 年，第 44 页。

[29] 《纽约时报》，1945 年 8 月 7 日。第 1 页。

[30] Henry D. Smyth，《原子能的军事用途：美国政府发展原子弹之官方报告，1940~1945 年》（*Atomic Energy for Military Purposes: The Official Report on the Development of the Atomic Bomb Under the Auspices of the United States Government, 1940~1945*，普林斯顿：普林斯顿大学出版社，1945 年）。

[31] “爱因斯坦：$E=mc^2$”（A. Einstein, $E=mc^2$），《科学画报》。

[32] Jeremy Bernstein，《发现科学》（*Science Observed*，纽约：基本图书出版公司，1982 年），第 310 页。

[33] 伯恩斯坦曾写出狂想家有以下几个特点。这些人坚持认为自己的工作解决了所有问题。他们非常严肃，认为旁人都是来窃取他们的思想的。他们大量使用大写字母，深信媒体会对他们的工作感兴趣。“科学狂想家”（Scientific Cranks），见《发现科学》（Science Observed），第 14 章。

8

金　蛋
爱因斯坦的广义相对论方程

$$G_{im} = -\kappa(T_{im} - {}^1\!/_2\, g_{im}T)$$

说明：时空决定了物质如何运动，物质决定了时空如何弯曲。
发现者：爱因斯坦。
发现时间：1915 年。

人类思想上最伟大的成就之一……

——J. J. 汤姆逊（J. J. Thomson）

从当时以至现在，在我看来，这个理论都是人类对自然的思考中最伟大的一个。它是哲学渗透、物理学直觉和数学技巧三者最令人叹为观止的结合。不过它与人们日常生活体验的联系较少。它就像一件伟大的艺术品，吸引我远远地去欣赏和崇拜。

——马克斯·玻恩

爱因斯坦的广义相对论场方程描述了时空的弯曲。它与表述质能互变的狭义相对论不同，不是一看就能明白的。不过它也得到了公众的很多关注。

那天是 1919 年 11 月 6 日，地点位于伦敦皇家学会大楼的一个房间。房间的内部很像小教堂。中间廊道的两旁有几排椅子，几根柱子作为墙壁的内衬。房间的后边还有一个接待室，人多的时候可以使用。墙上挂着皇家学会最著名的会员牛顿爵士的肖像。

这一次的事件是伦敦皇家学院和皇家天文学院会员的联合会议。与会的听众想要看看日食中收集到的数据报告。此次日食发生在 6 个月前的 5 月 29 日。日食期间，科学家们拍摄了恒星照片，看星光在经过太阳附近时是否发生了弯曲。新闻界从有些与会人员那里已经知道了此次日食的重要性。伦敦的《时代》杂志用了很长的篇幅，写道："在科学界，有史以来人们首次被激发出极大的兴趣，希望对基本物理问题的'对手理论'进行检验。参会的天文学家和物理学家为数众多。"[1]这些对手理论当中就有一个是爱因斯坦的。他提出的空间是"弯曲的"观点是他的广义相对论方程的一部分。这一观点意味着星光在经过太阳时会发生弯曲。电子的发现者、英国著名物理学家汤姆逊主持会议。他宣布："这是自牛顿以来，在有关万有引力理论的问题上取得的最重要的结果。"汤姆逊把这一结果说成是"人类思想上最伟大的成就之一"。[2]会上作为听众的哲学家阿尔弗莱德·诺斯·怀特海，后来写道：

> 会议的整个热烈氛围完全就像是一部希腊喜剧：我们就像合唱团，评论着极为崇高事件的发展所揭示出的命运的裁决。演出具有强烈的戏剧性，传统仪式和背景中的牛顿肖像提醒我们，两个多世纪后的今天，牛顿定律第一次接受修正，最伟大的科学发现就诞生于现在。这对个人而言也是有益的：思想上的探险最后终于安全靠岸。[3]

狭义相对论也是一次探险，不过它的参与者涉及科学体的众多成员，包括菲尔兹杰拉德、洛伦兹、庞加莱和其他一些对牛顿和麦克斯韦之间的矛盾感到困惑的人。广义相对论则不同，它几乎是爱因斯坦一人独自完成的。多年来，他对广义相对论的研究经历了一个迷宫般曲折的过程。在这一过程中，思路之门常在不经意间开启。不过，走进死胡同也是常有的事。爱因斯坦常常需要为此后退好几步，把数年的工作推翻重来。只有到达终点的时候，别人才意识到爱因斯坦所走过的戏剧之旅以及他的工作有着多么非凡的重要意义。

旅行独奏曲

广义相对论的产生最初也是源于一个假想实验。这个假想实验要追溯到 1907 年 11 月。爱因斯坦这样写道：

> 那时我在伯尔尼专利局工作。坐在椅子上时，突然一个想法闪过脑际：如果人做自由落体运动的话，那么他是感觉不到自己的重量的。我很惊讶。这一简单的想法给我留下了深刻的印象。它促使我去研究引力理论。[4]

爱因斯坦将这一想法描述为“一生中最令我快乐的思想”。[5]

这一假想实验是关于狭义相对论的假想实验的雄心勃勃的延伸。狭义相对论的假想实验涉及的是匀速运动的情形，关键之处在于无法判断物体是否在运动：举例来说，有轨电车中的所有物体的物理行为在电车处于运动或者静止状态时是完全相同的。但是，新假想实验涉及加速运动。想象将电车提到一定高度后释放，再落回到地面的情形。爱因斯坦发现无法区分电车是在加速场中下落，还是在没有任何引力场的真空中下落。如果放开一个物体，比如钥匙、皮球或硬币，在与之同步的参考系中来看，他们将仍旧保持静止状态。爱因斯坦把这称为“最令他高兴的思想”，因为他发现无法区分这两种情况这一点很重要。

如今，多亏有林林总总的新闻节目报道宇宙飞船和航天飞机中的失重条件，我们才不会像 28 岁的爱因斯坦那样吃惊地看待这一思想。通过“引力场中的加速落体与无力场条件下的加速落体无法区分”这一思想，爱因斯坦发现存在引力就是存在加速度，即力。情况也可以颠倒过来：如果电车在地面上，那么能够分清电车是位于引力场中，还是正被加速吗？这些思想集中了爱因斯坦的不满足感，使他开始研究“无法说出这其中的不同”究竟意味着什么。

这个假想实验与催生出狭义相对论的假想实验所涉及的不满足感是不同的。后者是在尝试将两个完整系统，牛顿体系和麦克斯韦体系，整合到一起时所产生的矛盾；而前者是由两个想象中完全不同的事物实际却相同而产生的。由于传统和习惯的关系，人们在处理不同的事物时，似乎已经习惯于与不同的对象打交道了。其实，我们可以发现这些对象的作用其实都是一样的，并考虑怎么会出现这样的事情。这两种不同的事物就是惯性质量和引力质量。根据牛顿的理论，引力是一种特殊的力。较重物体所受到的引力比较轻的物体要来得大。不过非常巧合的是，较重物体的惯性质量较大，对引力的抵抗程度也较大，其总的结果是恰好使得所有物体的加速度都相同。爱因斯坦想：让我们假定没有什么巧合，看看会发生些什么。

这件事的另一种说法是爱因斯坦在寻找另一种协变性，这种协变性比从早期相对论中已发现的协变性更加深刻。简单地说协变性就是所谓的客观性的一部分：说某个物体是世界的真实部分就等于是说该物体从不同的“角度”看起来是不同的——不仅包括照明条件，还包括了空间位置和速度等。这种不同角度下看到的不同可以用“变换”准确得出。因此，协变性似乎使物体的外观与物体的实质之间产生了不同。在 1905 年的理论中，爱因斯坦发现不管物体是否运动，得到同一描述的变换都是适用的。即便物体的速度已经接近光速，“看上去”已经出现了不同，只要物体匀速运动，变换就仍然适用。如果物体在以接近光速的速度运

动时，看上去并没有什么不同的话，就有理由说该物体不是实际物体，不是世界的一部分。如今，爱因斯坦要将协变性拓展到加速系统：在参考系处于加速运动的条件下，“真实”物体的描述将会发生什么变化？由此出发，爱因斯坦对他的工作进行了大胆的重新表述，称为“广义相对论”，以区别于它的前身“狭义相对论”。

第一步：等价原理（1907 年）

开辟了独奏曲旅程的假想实验是爱因斯坦在为《放射性年鉴》（*Jahrbuch der Radioaktivität*）写一篇总结相对论理论的论文时想到的。论文中，爱因斯坦在最后加进了一个 10 页的“新思想”。他写道：直到现在，他还是把相对论应用于匀速运动系统，但“是否可以认为相对论原理也适用于具有相对加速度的系统呢”？如果有两个系统，一个系统具有一定的加速度，另一个系统在施加了同样力的均一引力场中保持静止，那情况又是怎样的呢？爱因斯坦说，就目前所知，两个系统的物理定律是相同的。因此，应该假定引力场和参考系中相应的加速度在物理上是完全等价的。[6]爱因斯坦写道，他并不知道这一“等价原理”是否正确。他只想看看如果原理正确，将会得到什么结果。

接下来的几页中包含了许多日后出现的相对论的重要内容。他得出了惊人的结论，如位于引力场中高处的人将感受到低处时钟远离（源）的速度减慢，以及引力场会影响光的传播路径。尽管“很不幸，地球引力场的效应如此微弱，以致无法将理论结果与经验进行比对”，[7]然而爱因斯坦仍能获得可验证的预测结果。

例如，爱因斯坦开始关注悬而未决的水星轨道问题。19 世纪中叶，天文学家注意到水星绕太阳运行时在离太阳最近的点（叫做近日点）并不是固定不动的，而是绕着轨道缓慢移动。起初，天文学家认为这是由于受到其他行星的引力场的影响，可是当这些影响被一一考虑进来后，还有一个小量（每 100 年 43 秒的弧度）无法确定归属。这一偏差并不太

小，不能忽略；但也不是太大，不足以对牛顿系统构成怀疑。可在牛顿系统中，这的确又是个问题。于是有人设想可能存在看不见的伏尔坎（Vulcan）行星之类的天体，或者在太阳周围存在一层椭圆盘状的星云物质。不过人们未能发现它们的存在。另一种方式就是修正牛顿的理论，在平方反比定律中引入微小的变化。不过这种方法可能会带来不希望产生的副作用。半个世纪以来，水星轨道的进动问题一直是天文学上的一个未解之谜。

爱因斯坦认识到，进动问题最终可由一种新的第一原理得到解释。这多亏爱因斯坦的理论中有牛顿理论中所没有的一些项，于是科学家就又可以通过预测宇宙来认识未发现的现象，也无需为此调整常数和修正公式。爱因斯坦在 1907 年圣诞前夜给他的朋友哈比希特（Habicht）的一封信中写道："现在，我正在进行引力定律的相对论分析，希望我所采用的方式能够解释目前仍无法解释的水星近日点的长期变动问题。"[8]但是多年来，这些变动仍然是一个谜。

从 1907 年到 1911 年，爱因斯坦没有写过任何有关引力的东西。一方面这是因为个人生活上的一些影响。1910 年，爱因斯坦的第二个孩子出生；1911 年 3 月，他与妻子和两个孩子移居到了布拉格。爱因斯坦的新工作，不管是在瑞士联邦理工学院（Eidgenössische Technische Hochschule，ETH）做副教授，还是 1911 年在布拉格，都给他带来了更大的压力。此时爱因斯坦全身心地投入到量子论问题的研究中。在该过程中，他发现了如何"驱散"开尔文所谓的"第二朵乌云"：在将麦克斯韦–玻尔兹曼理论应用于某些实验结果时所遇到的困难被解释为量子物理效应。

在爱因斯坦回到了引力的研究之前，迷宫中有一扇新的希望之门打开了，尽管在当时爱因斯坦错过了这扇门。它是由爱因斯坦以前在 ETH 时的一位教授赫尔曼·闵可夫斯基（Hermann Minkowski，1864—1909 年）开启的。闵可夫斯基并不怎么喜欢爱因斯坦这个学生，有次还叫他

“懒虫”。闵可夫斯基看到 1905 年爱因斯坦发表的论文后，说道：“真是没想到那家伙能想出这么聪明的东西来。”[9]实际上，爱因斯坦才是闵可夫斯基最理想的学生，他借鉴、吸收并转化了自己的所学，为自己的老师提供了指导。

闵可夫斯基受爱因斯坦狭义相对论的启发所开启的这扇门用一种数学方法将空间和时间放到了相同的位置上。闵可夫斯基所描述的物体不仅具有 x、y 和 z 的位置坐标，而且还有与第 4 条轴，即时间轴，相对应的 t 坐标。也就是说，这一新的方式把物体看成是沿着时间线运动的，就像沿着空间中的线运动一样。原地不动的物体可以用一条直线表示，变化的只是物体在 t 轴上的位置。如果物体沿着 x 轴做匀速直线运动，那么只需要用一条曲线来表示它。因为物体此时不仅沿着 x 轴运动，同时也在沿着 t 轴匀速运动。运动情况更复杂的物体需要用更复杂的曲线表示。两个不同位置上的物体之间的距离需要用 4 个项表示，而不是 3 个。这就好比在毕达哥拉斯定理中加上了另一条与两事件之间的时间差相对应的“边”。三维欧几里得空间中的毕达哥拉斯定理是 $s^2=x^2+y^2+z^2$，其中 s^2 对所有观察者来说，长度都是不变的。现在，将这一定理推广到“空间–时间”中就变成了 $s^2=x^2+y^2+z^2-(ct)^2$。

闵可夫斯基公式进一步拓展了协变性，在客观实在的条件下引入了物体的时间行为。只用 x、y 和 z 位置坐标，而不用 t 坐标描述的物体，就好比是一个人的特写与录像之间的差别。没有 t 坐标的描述是片面的。在空间–时间中，物体的全面描述就更像是一部录像，因为它同时刻画了物体的空间位置和时间。为得到这些公式，闵可夫斯基采用了新的数学工具（现称为“张量”），把几组量从一个坐标系变换到另一个坐标系。张量是按指数划分成“阶”的数学对象，描述的是指数的复杂程度。0 阶张量是标量，1 阶张量是向量，2 阶张量是一组具有许多元素的复杂矩阵，可以将一组坐标变换为另一组坐标。闵可夫斯基采用 2 阶张量，把空间和时间结合了起来。在该过程中，他几乎与爱因斯坦同时对二者进

行了彻底的提炼。1908 年，闵可夫斯基在科隆做了一个报告，开头是这样的：

> 先生们，我要向各位展示的空间和时间的观点源自实验物理学，也应用于实验物理学。实验物理学是基础。这些观点是很大胆的。从今以后，空间、时间都注定要淡去、成为泡影。只有将二者结合起来，才能保存独立的实在。[10]

不过，此时爱因斯坦在迷宫中也经过了这扇门。可他没有把老师的工作放在心上，认为只不过是些“多余的学问”。[11]事实证明，爱因斯坦还是得回头。

1911 年，爱因斯坦与妻子米列娃和两个孩子一起住在布拉格。继 1907 年的论文“引力对光的传播的影响”（On the Influence of Gravitation on the Propagation of Light）之后，爱因斯坦又发表了一系列论文。[12]爱因斯坦开头写道：在一篇早期的论文中，我解决了引力是否对光产生影响的问题，但“之前对该课题的处理并不令我满意……因为我现在意识到，那种分析最重要的结果之一就是可以用实验来检验”。实际上检验有两种。根据等价原理，“不能再提及参考系的绝对加速度，就像不能在一般相对论中提及系统的绝对速度一样”。这对了解光在引力场中的行为是有意义的。爱因斯坦在第三部分作了以下讨论：第一个意义是引力红移的存在，即引力场（如太阳表面）中的光源所发射出的光与地球上类似的光源相比，会发生红移。测出这种不同对于验证广义相对论而言非常重要，尽管测定很困难。

第四部分讨论了第二个验证，即星光的弯曲。因为“经过太阳附近的光线受引力场的影响会发生折射，因此太阳附近的恒星与太阳之间的角距将明显增加，增加量将近有 1 秒”。假定科学家在日全食时拍下背景中恒星的照片，并将这些照片与没有太阳时同一恒星的照片加以对比。爱因斯坦怀疑，如果星光在经过太阳时发生了弯曲，那么照片将是不同

的，并且前一张照片里的恒星看上去离太阳较远。目前爱因斯坦已经预测出了距离到底有多远。

> 光线经过太阳时的弯曲量为 4×10^{-6}=0.83 秒弧度。发生日全食时，可以看到天空中太阳附近的固定恒星，故而可将理论结果与经验相对比。如果天文学家也来研究这里提出的问题，那就再好不过了。因为没有任何理论，也就无法知道现有设备能否探测出引力场对光的传播的影响。

这一弧度值（不久后修正为 0.87 秒）反映了这样一个事实：根据 $E=mc^2$，光也是有质量的。根据牛顿原理，光与石头和苹果一样会受到引力的作用。实际上，该值称为“牛顿值”。

在 1911 年的论文中，爱因斯坦对当时仍然是谜的水星进动问题一无所知。后来这篇论文却成了广义相对论的一个关键预测。前两个预测——特别是星光的弯曲问题——则产生了难以预想的结果，在 8 年后的一次著名会议上达到顶峰，并为世人所知。同时，爱因斯坦也在说服天文学家们参与到预测的研究中来。

1911 年 8 月，他将“星光的影响”（Influence of Starlight）一文送到了乌得勒支大学的一位天文学家那里。这位天文学家写过一篇关于太阳红移的论文。在文中，爱因斯坦提到了自己得到的“有点惊人的”结论——“引力势差”可能是造成太阳红移的原因。“引力场导致的光线弯曲也可从这些论证中得出。”爱因斯坦还补充道：如何还有其他什么原因的话，“那么我的宝贝理论就得扔进废纸篓了”。[13]同月，爱因斯坦将论文交给了欧文·弗罗因德里希（Erwin Freundlich）。弗罗因德里希是柏林的一位年轻的天文学家，后来成为在天文学家中宣传相对论的人。他答应看一下木星对星光的影响，并对日食期间拍摄的照片进行研究。他证实木星还不够大，无法使星光弯曲任意可探测到的

角度。得知这一消息的爱因斯坦叹道："要是有一个比木星大一点的行星就好了！""可是大自然觉得让人类轻而易举地发现她的规律并不是自己的分内之事。"[14]弗罗因德里希也开始研究能否由以前发生日食时拍摄的照片探测到恒星折射，但结果以失败而告终。他启动了一个考察计划——拍摄 1912 年 10 月将于巴西出现的日食。不过这次考察因为下雨被迫取消。有位组织者挖苦地讽刺道："我们没发现日全食，却深受日全食之苦。"[15]

弗罗因德里希开始计划到俄国考察，观察将于 1914 年 8 月 21 日发生的日食。1913 年 6 月，爱因斯坦激动地写信给欧内斯特·马赫（Ernst Mach）："明年日食发生时，我们就能知道光线是否会受到太阳的影响而发生弯曲。换句话说，也就会知道参考系的加速度与引力场是等价的这一基本假设是否正确。"[16]同一年秋天，爱因斯坦还写信给天文学家乔治·海尔（George Hale），问有没有一种足够强大的望远镜，能够在白天看到太阳附近的恒星。如果有这种望远镜的话，就无需一定要等到日全食了。不过这位著名的天文学家粉碎了他的希望，所以爱因斯坦只能等 1914 年 8 月发生的日食了。可惜天不作美，考察也遭遇了灾难。灾难的发生有政治上的原因，也有天气的原因。6 月 28 日法朗西斯·费迪南大公（Archduke Francis Ferdinand）遇刺，随后奥匈帝国就侵犯塞尔维亚，同时德国也于 8 月 1 日宣布对俄战争。弗罗因德里希的探险是战争早期的牺牲品。他本人也被捕，虽然不久就被赎出，可以返回柏林，但设备被悉数没收。其他的几次探险也因为天气原因而搁浅。

爱因斯坦尽其所能地提供帮助。可面对挫败，他也有些耐不住性子了。不过从长远来看，探险尝试的失败对爱因斯坦还是有益的。在俄国探险之时，爱因斯坦已经进入了迷宫的另一个不同的部分，这促成了他对预测的修改。

第二步：时空几何学

同时，从某种程度上来说，多亏研究了光在引力场中的弯曲问题，爱因斯坦才认识到等价原理更深层的意义——方程必须对所有类型的几何空间都成立。具有几何学的空间并不意味着空间是弯曲的（只有与被认为是直的事物比较而言，才可以说某个事物是弯曲的），而是与通过空间的物体（如一束光）的路径测量和有关。如果引力场造成了光线的弯曲，那么光线路径的测量就会按某种几何学的方式加起来。这其中涉及“不把引力看作力（即施加拉力的事物），而把它看作空间本身的性质”的思想。通过空间的物体都需要遵循这一几何结构。爱因斯坦写道：如果写出的方程中不涉及空间几何结构，那么方程就好像是“不用字句就来描述思想”一样。[17]而且，他也认识到了自己以前的老师闵可夫斯基（时已去世）工作的价值，以及张量的采用对目前工作所带来的极大简化作用。于是，爱因斯坦在迷宫中折回了几步，进入到闵可夫斯基开启的那扇门中。可此时的他却发现自己不知所措，无法找到适用于非欧几里得弯曲几何学的张量。

幸运的是，爱因斯坦知道到哪里寻求帮助。1912 年夏天，他离开布拉格，来到苏黎世。既是他以前在苏黎世的同事，也是同班同学的马塞尔·格罗斯曼（Marcel Grossmann）为他提供了一个职位。两位同窗昔日一起学习的时候，爱因斯坦常常借格罗斯曼的课堂笔记，这才得以一心用在物理学上。如今，格罗斯曼已是 ETH 数学–物理分院的院长。“格罗斯曼，你必须帮我，不然我就疯了！”爱因斯坦绝望地写道。[18]于是，格罗斯曼向爱因斯坦介绍了数学家发明的非欧几何，并向他解释张量微积分。非欧几何是由波恩哈德·黎曼（Bernhard Riemann，他提出了一个有 20 个元素的曲率张量）、格雷戈里奥·里奇–库尔巴斯托罗（Gregorio Ricci-Curbastro，他提出了一个只有 10 个元素的简化版本）和图里奥·李微–西威塔（Tullio Levi-Civita）共同建立起来的。在格罗斯曼的帮助下，

爱因斯坦开始着手提出一个适用于四维时空几何的一般的、闵可夫斯基式的张量，以表示引力场。8 月，他在给一位朋友的信中激动地写道："有关引力的工作进展得很顺利。我现在已经发现了最一般化的方程。如若不然，那就是我彻头彻尾地错了。"[19]10 月，他又写信给另一个朋友：

> 我现在正专心研究引力问题。相信在这儿的数学家朋友的帮助下，我可以克服任何困难。但有一件事是肯定的：我一生中还从未因为任何事情把自己搞成这样。现在，我对数学充满敬意。我觉得它细微的部分真是一种奢华，尽管现在我仍对这些部分一无所知。与这一问题相比，原来的相对论简直就是儿戏。[20]

第二年，也就是 1913 年，爱因斯坦和格罗斯曼联合发表了一篇论文，题目是《广义相对论和引力理论提要》（*Outline of a General Theory of Relativity and a Theory of Gravitation*）。这篇论文离最终方程仅有"一步之遥"。[21]它包含两部分，一部分是物理学，由爱因斯坦写成；另一部分是数学，由格罗斯曼写成。在写作的过程中，他们发现有迹象表明广义协变性是不可能的。这一发现极大地困扰了爱因斯坦，他将其称为理论中的一个"肮脏的黑点"。[22]他放弃了场方程广义协变性的希望，转而研究其他，却又走进了死胡同。最后，爱因斯坦用了两年时间才从中解脱出来。

1919 年初，爱因斯坦离开了苏黎世。他在柏林找到了一个新的职位。同爱因斯坦分开的除了格罗斯曼，还有他的妻子米列娃，原因是两人的婚姻破裂。在 ETH 的一个老朋友为他举行的告别宴会上，爱因斯坦坦陈了自己的担心，抱怨自己被人当做"一流的母鸡"来对待。他告诉送他回家的人说："其实连我都不知道自己还会不会下蛋。"[23]

第三步：协变方程

爱因斯坦一到柏林就开始全神贯注地研究广义相对论。有一天，弗罗因德里希来到他的书房，发现天花板上挂着一个钩子，上面挂着没时间读的信。[24]1915 年 10 月到 11 月间，爱因斯坦发现了自己的错误，回过头去寻找一个具有广义协变性的公式。在给朋友的一封信中，爱因斯坦写道：那“将是我一生中最刺激、最精疲力竭的时刻”。[25]11 月 4 日，爱因斯坦告知普鲁士科学院，他已对自己之前报道的“场方程失去信心”。他解释说自己已经回到对场方程的“广义协变性的要求”上。“三年前，虽然心情很沉重，我还是放弃了这一要求”，并提出了一个真正的协变理论。“人们只要抓住了它，就一定会被它的魅力所吸引。”

接下来两个月里的一天，爱因斯坦作出了一个发现。一位传记作家写道：“这是爱因斯坦科学生涯，甚至可能是爱因斯坦一生中最强烈的情感体验。”[26]他发现新理论可以完美地解释水星近日点的进动问题。在这个他 1907 年所作出的预测上，爱因斯坦倾注了心血。现在，他的新理论无需任何假设或假定就可以解释这一预测。他在给洛仑兹的信中写道：“好几天我都难以抑制住自己的兴奋和激动心情。”[27]他告诉一位朋友说自己甚至出现了心悸，又告诉另一位朋友说自己的心情始终不能平静下来。[28]

11 月 18 日，爱因斯坦向普鲁士学会的会员们宣布了他的定律——能够对水星轨道作出解释的协变引力场方程。他同时还宣布了首次发现的另一个现象，即弯曲时空理论可预测出太阳附近星光的偏角是之前预测大小的两倍。“但是，从该理论出发得出的引力场对于光线的影响与从我之前的工作中所得出的有所不同。掠过太阳表面的光线的偏角弧度为 1.7 秒，而不是 0.85 秒。”旧的牛顿值是 0.85 秒，而最新的“爱因斯坦值”是 1.7 秒。[29]

11 月 25 日，在一个题为“引力场方程”（*The Field Equations of*

Gravitation）的报告中，爱因斯坦写出了如下我们今天所熟悉的方程：[30]

$$G_{im} = -\kappa(T_{im} - \frac{1}{2} g_{im} T)$$

不过，有些情况下人们会把 T_s 写成 R_s，并加进其他项。[31]

方程有两部分。左边是表达空间几何的一组项。右边是描述能量和动量分布的一组项。左边涉及的是几何，右边涉及的是物质。物理学家约翰·惠勒（John Wheeler）这样描述：从左到右读是时间-空间告诉能量如何运动，而从右到左读则是质量告诉时间-空间如何弯曲。虽然方程仅仅描述了人们所认识的世界中的实验上微不足道的偏差—— 一些离位恒星的位置，但是它却是牛顿世界概念上的革命。在这个新世界中，不存在绝对时间和绝对空间，引力也不再是力——不是物体之间的拉力——而是空间和时间的性质。

爱因斯坦对理论的正确性非常自信。他给物理学家阿诺德·索末菲（Arnold Sommerfeld）送去一张明信片："研究过广义相对论之后，就会确信它是正确的。因此，我不想为它做任何辩护。"[32]然而，虽然爱因斯坦本人对广义相对论的核心了然于心，但是对其他几乎所有的人来说，广义相对论仍然像一座迷宫，深不可测。爱因斯坦在给洛伦兹的信中写道："基本公式很好，但是推导出来的公式就比较糟糕了。"[33]于是在1916 年初，爱因斯坦开始坐下来构建一条逻辑路线，让其他人也能够理解。随后，他在 3 月的一期《物理学年鉴》上发表了一篇长达 50 页的论文，题目是"广义相对论的基础"（The Foundation of the General Theory of Relativity）。这篇论文获得了极大的成功，以小册子的形式重印多次，并被译成英语。论文的最后一部分陈述了由理论得出的 3 个实验预测：[34]

> 从大的恒星表面到达地球的光谱线看上去移向光谱中红色的一端。
>
> 经过太阳的光线将产生 1.7 秒弧度的偏角，经过木星的光线将产生 0.02 秒弧度的偏角。

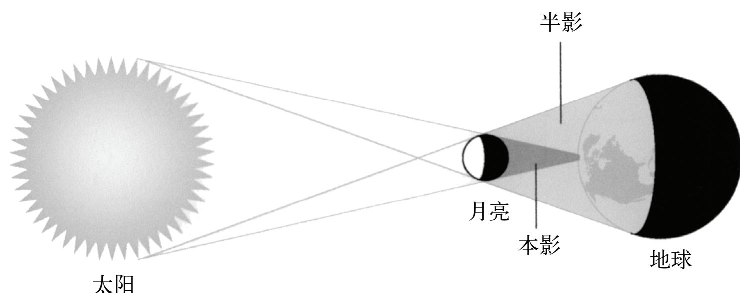
行星的椭圆轨道缓慢地发生转动。计算结果表明，水星的轨道每100年转过43秒，与天文学观测结果完全相符（勒维耶）。

第一个预测在当时很难验证，第三个在初次出现时已经包含在了理论中。但第二个（必须是经过太阳的光线）看上去好像是可以验证的。

大自然没能给人们提供多少帮助。

实验科学是一门利用已经理解的事物去探索未知事物的艺术。让球沿着斜面滚下并对其计时，对钟摆的摆动进行计时，测定油滴在电场中的行为等都属于实验科学的范畴。真正了不起的是（也即科学的奇迹）收获的总是比投入的要多。在某些情况下，可以在实验室中把这些事件筹划成“御前表演”（command performance），完全置于自己的控制之下。

日食就是这些伟大的宇宙表演中的一个。



宇宙表演

从距离地球和月球很远的地方用高倍天文望远镜看去，两者浸没在太阳发射的光之中。地球和月球各自绕着对方旋转，并在对方表面投射出锥形的阴影。人们根据牛顿运动定律，可以对这类运动准确地加以预测。地球和月球在转动时，常常会进入对方所投射的阴影区域当中——有时候是全部进入，有时候是部分进入。令人不可思议的是，有时在地球上看来，月球的大小与太阳的大小恰好相等，于是月球会完全挡住太

阳光，将阴影区内的物体置于一片黑暗之中，从而在白天时遮挡住恒星的光。由此形成的日食就为验证星光在经过太阳时是否会发生弯曲创造了条件。

爱因斯坦对验证自己的理论一事既担心又激动。换作别人，心情也是一样的。这不仅仅是一个公式的问题——它指出了宇宙的基本结构不同于人们的已有认识，而且比人们一直以来所认为的要更加奇怪。对于前几次探险的失利，爱因斯坦很是沮丧。不过这也避免了爱因斯坦陷入对日后尚需修改的不成熟结果进行验证所带来的窘境。

在概述包含三个预测的里程碑式的论文“广义相对论”于 1916 年发表之后，爱因斯坦再次开始积极推动日食期间的理论验证工作。他将论文的副本寄给了荷兰天文学家和物理学家威廉·德西特（Willem de Sitter）。随后，德西特又将副本送到了皇家天文学会的秘书亚瑟·爱丁顿（Arthur Eddington）那里。由于战争中断了通信，因此这是英国仅有的一份爱因斯坦论文的副本。身为物理学家的爱丁顿一直对水星的偏差甚为困惑。他对爱因斯坦的理论很感兴趣，之后还写了几篇这方面的综述文章。而且，作为一名贵格会教徒（Quaker）与和平主义者，爱丁顿也很想借此化解德语国家科学家和英语国家科学家之间的敌对情绪。但是，他的努力遭遇到了阻力。英国物理学家奥利弗·洛奇（Oliver Lodge）几乎与赫兹同时发现了电磁波的一种形式，而且他坚信以太是存在的，称以太漂移可以解释水星的进动。还有一个叫作托马斯·西伊（Thomas J. J. See）的美国人也反对爱因斯坦的观点。他坚持认为引力是一种实际存在的物理力，并写道：“整个相对性学说都是建立在错误的基础之上的。将来总有一天，人们会发现相对性其实是无源之水、无本之木。”[35]但是，爱丁顿坚信爱因斯坦是正确的，并为日食探险募集资金，来验证这一迷人的基本理论。[36]最近有些社会建构工作者开始注意到爱丁顿在使科学界、媒体和公众关注相对论这一过程中所起到的作用。这些人把全局利益也看成是个人私利，暗指爱丁顿实际上所做的不过是操作、公

共关系和自我推销。[37]但是，爱丁顿对这一工作感到兴奋是再自然不过的事，他希望与他人分享自己的激动心情。同样地，公众和媒体积极响应这一具有潜在重大意义的消息也是很自然的。前面那些人的想法暴露了他们对科学家和公众的幼稚认识。假若爱丁顿虽然对这一预示着人类对世界认识的革命性变化的科学工作怀着激动心情，但却不想与人共享，那他可就真是有问题了。

最后一次日食考察的遭遇并不比之前的考察好到哪里去。由于战争原因，原定于 1916 年的委内瑞拉日食观测探险活动搁浅了。1918 年，机会再次来临。这一年，美国发生了一次日食，但是因为天气恶劣、加利福尼亚大学 Lick 天文台借给 1914 俄国探险队的设备尚未归还以及组内成员之间存在矛盾等原因，最终考察结果未能发表。

爱丁顿协助邀请到了英国皇家天文学家弗兰克·戴森（Frank Dyson）。戴森是提出要对 1919 年 5 月 29 日即将出现的日全食进行观测的第一人。这次日食将从巴西北部，越过大西洋，穿越非洲大陆一直到非洲北部。这次日食将在数颗明亮恒星的背景下发生。1917 年 11 月，皇家学会的联合常设日食委员会和皇家天文学会组织了两次分别前往两个不同地点的考察，去拍摄日食照片。克罗姆林（A. C. D. Crommelin）和戴维森（C. R. Davidson）带领一队，到达巴西东北部的索布拉尔。另一队是去往普林西比岛。该岛属葡萄牙，长 10 英里，宽 4 英里，距离非洲西海岸有 120 英里。爱丁顿也参加了这一队的考察。但是，日食可以预测，天气却没法预测。索布拉尔的天气尤其令人担忧，5 月恰好是雨季的最后一个月。

1919 年 5 月 29 日的日食不过是一次普通的日食，也是地球再次进入到月球的锥形阴影区域的时间。不过这次日食后来却成为科学史上最重要的一次日食。

1919 年 5 月，两个考察队离开英国的格林尼治，登上蒸汽船安塞姆号，在葡萄牙的里斯本短暂停留之后，到达了马德拉。之后分成了两个组。索布拉尔小组继续乘坐安塞姆号去往巴西，另一小组继续在马德拉

停留 4 周时间，等候去往普林西比岛的轮船。

索布拉尔考察小组到达距离沿海 80 英里的相对荒凉的内陆之后，在确定日食发生前没有比赛安排的前提下，开始在当地的赛马俱乐部的赛马场跑道上搭设仪器设备。巴西政府安排了搬运工、砖匠、木匠和翻译。另外还从里约热内卢购买了一辆汽车供考察使用（这也是在索布拉尔见到的第一辆车）。为了避风，考察队为两台望远镜搭建了支撑结构。其后突然来了一阵旋风，掀翻了支撑结构。不过木匠们很快就从别的地方找来了梁，对结构进行了修复。5 月 25 日大雨倾盆，提示考察队员们现在尚是雨季。更添乱的是两个望远镜中较大的那个的驱动装置运行不正常，而且两台望远镜都有聚焦问题。5 月 29 日，月亮的影子开始横扫地球表面。当天早上，索布拉尔的考察队员醒来发现是阴天。日食开始时，太阳还躲在云层后面。但在出现日全食的前一分钟，太阳跳出了云层。在短短的 6 分钟内，克罗姆林和戴姆森不停地拍摄照片，每 5~6 秒就曝光一张底片。“日食很棒。”他们在发回的电报中说道。

另一考察队在普林西比岛登陆后，在西北部的大农场搭了一个棚子。时常出现的阴天令队员们不安起来。5 月 29 日，队员们醒来发现大雨倾盆、雷电交加，完全看不到太阳。在出现日全食前半小时，队员们看到了彼时呈弯月形的太阳，才增添了些信心。但是乌云始终未能褪去，爱丁顿和同事科廷汉只能无助地抱着最后一线希望，拍些乌云和短短一瞥的恒星。“希望透过乌云能得到点有用的信息吧。”他们在发回的电报中写道。

接下来的几个月里，两个小组在考虑了各种可能误差来源的基础上，开始计算光线的偏折量。

“今天的好消息”

9 月，爱因斯坦焦虑地写了一封信给洛伦兹，问他是否已听说英国人的结果。9 月 27 日，洛伦兹发回了电报：

爱丁顿在太阳外缘发现了恒星位移，初步测量结果显示位移大小在 0.9 秒到 1.8 秒之间。致以美好的祝愿。洛伦兹

爱因斯坦接到电报很快就给母亲玻琳送去了一个明信片。她病得很重，仅能维持几个月的生命：

亲爱的母亲，今天有好消息。洛伦兹给我来了电报，说英国人的探险已经表明了太阳附近光线的弯曲。[38]

爱因斯坦还向《自然科学》（*Naturwissenschaften*）发去了一封短信，表达自己的激动心情。[39]不过现在焦虑已经涣然冰释，他可以镇定下来，更加沉着地应对一切。爱因斯坦的学生伊森・罗森塔尔–施奈德（Ilse Rosenthal-Schneider）来拜访他的时候，他把电报拿给她看，说："看这个，或许你会感兴趣。"她后来回忆：

他给我看的是爱丁顿的电报。电报的内容是日食探险的测量结果。当我高兴地说结果与他的计算正好吻合的时候，他却很平静地说："可我早知道理论是正确的。"当我问如果没有事实能够证明他的预测怎么办时，他答道："那我可得为亲爱的主感到遗憾啦——理论就是正确的。"[40]

1919 年 11 月 6 日，伦敦皇家学会和皇家天文学会的联合会议由汤姆逊爵士主持。25 年前，汤姆逊发现了电子。他在开幕词上说："我提议请皇家天文学家向大家报告一下五月日食探险的结果。"[41]

戴森讲道：

探险的目的是为了确定太阳的引力场是否会引起光线的位移；如果答案是肯定的，那么位移的大小是多少。爱因斯坦的理论预测位移的大小与光线离太阳中心的距离成反比。对掠过

> 太阳的恒星来说，位移可达 1.75 秒。爱因斯坦的理论或引力定律已经解释了水星的进动问题。这一直是天体力学领域的一个突出问题，人们希望能藉此对爱因斯坦的理论进一步加以验证。

戴森又继续说，任何弯曲都会产生“将恒星抛离太阳”的效应——使恒星看上去离太阳更远。他简要回顾了探险中发生的事件，提到了在较大设备上所发现的不足。戴森是一位优秀的科学家，他知道数据质量并不是等同的，有些仪器的数据会好一些。因此不能单独对结果进行分析评价。他侧重于小设备的测量结果，得出结论：“已经得出确切的结论，表明光是按照爱因斯坦的引力定律发生偏折的。” 克罗姆林随后也认可了这一观点，并针对较大设备的不足给出了一个简明的解释。

爱丁顿的能力有点不如戴森，数据质量也不如克罗姆林。回忆探险的历程，他把失败归结于恶劣的天气。他竭尽全力从正面去描述自己的结果，去劣存优，指出普林西比岛的乌云和均一的温度有利于减少面镜失真。而面镜失真已经对大索布拉尔镜产生了负面的影响。爱丁顿坦言自己“尽可能充分地利用仅有的一小点材料”。他发现太阳附近 1.6 秒的位移值“支持在索布拉尔拍摄的照片”。他抛弃了太阳周围存在星云折射物质的解释，得出结论——结果支持爱因斯坦的定律，即光线弯曲的描述。这一定律未必可以完全确定爱因斯坦理论，不过它却是时空弯曲背后的思想。

汤姆逊又发话了。他把报告称作“重大的交流”，并说：“这是自牛顿时代（当时附近挂着的这位早期会员的肖像）以来有关引力理论获得的最重要的结果。爱因斯坦与学会一直都有紧密联系，应该举办一次会议，宣布这一结果。”

有些人认为理论的数学过于复杂，似乎超出了物理学家的能力范围。其中有一位说：“我不能想象一个重大的物理学事实无法用简单的语言表达出来。爱丁顿教授可不可以把他推崇的这篇论述由张量记号改写成

别的更简单的形式呢？”一位对此持有怀疑态度者，援引光谱移动一直不存在的这一事实，指出光线偏折这一事实（“一个独立的事实”）并不一定能确证爱因斯坦的理论。他指着牛顿的肖像说：“我们在修正牛顿引力定律的时候，一定要仔细。没有牛顿引力定律，还谈何修正？”然而，在不到两年的时间里，反对爱因斯坦的各种说法就都不攻自破了。

爱因斯坦的理论与牛顿的理论在实验结果上的不同是很小的——不过是某些恒星和光谱线位置的微小移动和水星轨道的微小晃动。但是这些不同也几乎不能再大了。它们表明宇宙构建方式与之前人们的理解存在根本上的不同。

在牛顿理论中，引力包含了吸引力（一种拉力）。所有物体都会对一定距离之外的其他物体施加这种吸引力。它瞬时发挥作用，并且处处存在，与物体之间的距离的平方成反比。受引力作用的物体会向着施加引力的物体加速运动。不同质量物体的加速度是相同的：物体质量越小，拉力越小；质量越大，拉力也越大。相反，在爱因斯坦的理论中，引力涉及的实际上是空间的弯曲。这种弯曲是由周围的物体的质量造成的。物质和能量在经过空间时是沿着对二者敞开的路径的。

引力定律是科学史上对基本概念的一次伟大的革命。正如爱丁顿所写的那样：

> 250 年后依然如故的牛顿框架确实是太粗糙了，无法与人们正在获得的新的观测结果相吻合。在没有更好的框架时，牛顿框架依旧被采用，但是其中的定义却被延伸到之前从未用过的场合下。而我们就像是图书管理员，仍按着 100 年前的学科体系对图书进行分类整理，并且还想为好莱坞、空军和侦探小说一类的图书找到合适的位置。[42]

1921 年，唯一一个仍对理论感到失望的大人物就是爱因斯坦自己了。他挑剔的眼睛想要在整个理论中看到对称性。现在的理论左边是坚实的

（因为是用时空几何表达的），右边则不是。他曾把这一理论比作规划不周的建筑，一半是“优质大理石”，另一半是“劣质木材”。[43]因为这种不满足感的存在，爱因斯坦在一生余下的大部分时间里都努力修复着这座建筑，结果却被证明是徒劳的。他在这上头花费了三十多年的时间，从未获得成功。



茶　歇

科学评论家

> 引力场方程把空间弯曲与物质分布联系在一起。现在这已是人尽皆知的事实。
>
> ——伊塔洛·卡尔维诺（Italo Calvino），
>
> 《宇宙连环画》（*Cosmicomics*）

> 于是，这个可怜的悖论，即爱因斯坦的发现（有史以来人类理性的最大胜利）被绝大多数人毫无怀疑地接受了。
>
> ——《时代》，1946 年 7 月 1 日

公众理解科学新进展的过程常常是以所谓的摩西和亚伦模型的形式出现的。科学家就像是摩西，一脚踏在神界的半球上，常把发现带到世间来。这种原始的活动通过某个亚伦传达，用图像和通俗语言翻译出来，帮助公众理解。

摩西和亚伦模型实际上是一个两步过程。这一模型时刻都在上演——例如在晚间电视新闻中，新闻发言人总是尽力在规定的 60 秒时间

里解释最新的进展。相比其他一些发言人，有些“亚伦”的发言更有效、更吸引人。

爱因斯坦的广义相对论给未来的亚伦带来了一种负担。要一步或两步就完成相对论的分析过程是很难的。这一理论涉及复杂的数学和人们不熟悉的思考方式。即使是物理学家，也得需要几年才能掌握。学习该理论就像是适应文化，没有捷径可走。有时候，新闻记者干脆举起双手，说没办法向非科学家解释这一发现。《纽约时报》在 1919 年 11 月 6 日公告发布之后就是这么处理的。物理学家赫尔曼 • 邦迪（Hermann Bondi）曾说，如果公众没有相对论玩具可以玩，那就无法理解相对论。

但是向外人解说科学就像是同非城市居民谈论城市一样——你得根据听众的反应决定谈话内容。如果他们想成为城市居民，你的谈话就要集中到规章、制度和法律等内容上，而且听众需要一定时间才能掌握。如果你的听众就是来旅游的，没有想要成为城市的居民，那你就要集中到公共旅游景点上，不要说得太细，很多完全可以省去。

为了使旅游者式的听众能够理解相对论，可以进行很多有趣的尝试。有一个方法就是选取并制作一些巧妙的图例，以易懂的方式说明含义：例如在双生子佯谬中，一个双生子在空间中以接近光速飞行，他衰老的速度不同于地球上的另一个双生子；再如宇航员不知道自己是否是在加速，以及是否位于引力场中。另一种就是采用传记的方式：读者可以看一下非常流行的沃尔顿 • 萨克森（Walter Issacson）最近的一本书《爱因斯坦的生活与宇宙》（*Einstein: His Life and Universe*），或者阿伯拉罕 • 派斯的《上帝是微妙的：爱因斯坦的科学与生平》（*Subtle is the Lord: The Science and the Life of Albert Einstein*）。第二本书的难度较大，获得过美国书卷奖（American Book Award）。另一种向外人传达广义相对论含义的方法是采用吸引人的照片，如橡胶板上的重物。重物会使橡胶板发生弯曲。弯曲量由物体的重量决定，并影响经过此处的弹球的路径——这与物体使时空弯曲是类似的，弯曲量由物体的质量决定，并影响经过此处

的物体（包括光）的路径。有些作者会同时采用三种方法——布莱恩·格林（Brian Greene）就在两本精彩的书《宇宙的构造》（*The Fabric of the Cosmos*）和《宇宙的琴弦》（*The Elegant Universe*）中采用过。

其他一些向外人传达复杂科学的巧妙例子还有埃德温·艾勃特（Edwin A. Abbott）的《二维国》（*Flatland: A Romance of Many Dimensions*）。这是一本著名的小说，内容是正方形与球体之间为理解多维空间进行的一系列对话。迈克尔·弗雷恩（Michael Frayn）的著名剧本《哥本哈根》（*Copenhagen*）则把玻尔和海森堡之间的冲突戏剧化，阐述了量子物理学中的许多问题。

不过，在当代科学讲述方面最需要强化的是所谓的科学评论。至少已有两人在倡导科学评论。一位是政治科学家兰登·温纳（Langdon Winner），另一位是哲学家唐·伊德（Don Ihde）。温纳指出：在艺术领域中，人们凭直觉把评论理解为“在艺术和听众之间搭起的一座桥梁”，“它发挥着重要的作用”。举例来说，文学评论家“检查文字，分析文学的特色，评价质量，并寻求更深层次上的理解，从而可能对读到同样文字的其他读者有所帮助”。不幸的是，温纳叹道，在科学上却没有出现过同类的活动。障碍之一是科学家倾向于把扮演评论家角色的人看成是嫌疑犯，好像科学评论天生就是反对科学或者坚持己见似的。

同时，唐·伊德也在积极呼吁科学评论，甚至已经提出了他所认为的科学评论的轮廓。“科学评论家必须保持信息的畅通，而不是简单地获取信息。评论家是业余人士，他们热爱这个主题，但又不是业内人士”。这就像艺术评论家并非真正的艺术家或者文学作者一样。科学评论家为什么一定不能是完完全全的业内人士呢？伊德的答案是：“人们对于自我评论这件事恐怕是做得最差的。”

按照温纳和伊德的说法，科学评论家应该具备一个基本的功能——他们应该对科学技术给政界（温纳）和人类体验（伊德）带来的影响进行评估。例如，温纳写出了科技制品的“政治”，伊德则写出了工具器械

给人类体验带来的变化。简而言之，温纳和伊德所提倡和实践的这种评论判断出了科学技术在社会中的存在性，具有明确清晰的道德和政治意义。

还有另一个模型——科学评论的补充模型，包含了另一种类型的解读，即概述科学发现对人类理解自身、世界和自身在世界中的位置所产生的影响。该模型不要求单步解读过程，但要求具有艺术评论所充当的多重类型的角色。它牵涉的科学评论的类型与艺术评论同样详尽和广泛。艺术评论的存在对于艺术文化的繁荣而言是必需的。这些必需步骤包括几种不同类型的作品构成的复杂领域。这些作品可以是书籍、文章和专栏，还可以是小说和戏剧及其注释和评论等。如此一来，科学产生的知识就可以存在于文化中，而不仅仅是存在于器械方面，而文化也可以通过这一过程加强自身。

这种模型可以称为*阻抗匹配*。在声学和电气工程中，阻抗匹配先提取信号（产生自扬声器），然后将其置于具有不同"负载"的新环境中——周围环境。置入的方式应保证信号可被探测到。该过程既不是单步的，也不是双步的。它要求光滑连续地匹配负载，并使负载逐渐减小。也就是说，科学论述有一个（大）负载，而公众语言具有另一个非常不同的负载。仅凭基础教育和科学普及并不能将这二者联系起来。实现这一点需要许多不同的重叠步骤。每一步所需要的都不仅仅是专业理论知识的问题，而是需要将公共问题和公众期望与信号联系起来。

我们为什么要费力探讨这个问题呢？系统现在的工作情况相对已经很好了。为什么在对它们进行解释说明之外，还要努力去找出广义相对论等科学工作在道德和精神上对世界的影响呢？一部分原因是避免盛气凌人，好像把公众当成小孩。爱因斯坦的广义相对论是代表了人类在理解世界基本原理方面最杰出的工作，因此公众最好也能了解一点广义相对论的知识，获得一些它对人们认识世界的意义的感觉，并以此为自己的义务。说得严重一些，这也是人类在寻求理解自身和自身在世界中的位置的一部分。最要紧的还是人类自身的人性。

注　释

[1] 《泰晤士报》（Times of London），1919 年 11 月 8 日，第 1 页。
[2] 引自阿伯拉罕 • 派斯《上帝是微妙的：爱因斯坦的科学与生平》（纽约：牛津大学出版社，1982 年），第 124 页。
[3] 怀特海，《科学与现代世界》（*Science and Modern World*，纽约：麦克米兰，1954 年），第 13 页。
[4] 引自派斯，《上帝是微妙的》，第 179 页。
[5] 同上，第 178 页。
[6] 爱因斯坦，《爱因斯坦著作集》（*The Collected Works of Albert Einstein*，普林斯顿：普林斯顿大学出版社，1989 年）第 2 卷，A. Beck 译，第 301~302 页。
[7] 同上，第 310 页。
[8] 1907 年 12 月 24 日爱因斯坦致哈比希特的信，见《爱因斯坦著作集》，第 5 卷，第 47 页。
[9] 引自 Ronald W. Clark，《爱因斯坦的一生与时代》（*Einstein: The Life and Times*，纽约：世界出版公司，1971 年），第 120 页。
[10] Emilio Segre，《从 X 射线到夸克》（*From X-rays to Quarks*，纽约：多佛出版社，1980 年），第 85 页。
[11] 引自派斯，《上帝是微妙的》，第 152 页。
[12] 爱因斯坦，“论万有引力对光的传播的影响”（*On the Influence of Gravitation on the Propagation of Light*），《物理学年鉴》（*Annalen der Physik* 35，1911），第 898~908 页，收于《爱因斯坦著作集》，第 3 卷，第 379 页。
[13] 1911 年 8 月 24 日爱因斯坦致威廉 • 朱利叶斯（Willem Julius）的信，见《爱因斯坦著作集》，第 5 卷，第 199 页。
[14] 1911 年 9 月 1 日爱因斯坦致弗罗因德里希的信，出处同上，第 202 页。
[15] J. Earman 和 C. Glymour，“相对论与日食：1919 年英国日食探险和之前的探险活动”（Relativity and Eclipses: The British Eclipse Expeditions of 1919 and Their Predecessors），《物理学史研究》（*Historical Studies in the Physical Sciences* 11，1980），第 61 页。
[16] 1913 年 6 月 25 日爱因斯坦致马赫的信，见《爱因斯坦著作集》，第 5 卷，第 340 页。
[17] 引自派斯，《上帝是微妙的》，第 311 页。
[18] 同上，第 212 页。

[19] 1912 年 8 月 16 日爱因斯坦致 L. Hopf 的信，见《爱因斯坦著作集》，第 5 卷，第 321 页。

[20] 1912 年 10 月 29 日爱因斯坦致 A. Sommerfeld 的信，见《爱因斯坦著作集》，第 5 卷，第 324 页。

[21] 《数学和物理杂志》（*Zeitschrift für Angewandte Mathematik und Physik* 62，1913），第 225~261 页。"头发丝的宽度"这一说法来自 John Norton，"爱因斯坦是如何发现场方程的：1912~1915 年"（How Einstein Found His Field Equations: 1912~1915），《物理学的历史研究》（*Historical Studies in the Physical Sciences* 14:2，1984），第 253~316 页。

[22] 1913 年 8 月 16 日爱因斯坦致 H. Lorentz 的信，见《爱因斯坦著作集》，第 5 卷，第 352 页。

[23] 引自 Clark，《爱因斯坦》（*Einstein*），第 173 页。

[24] 同上，第 199 页。

[25] 1915 年 11 月 28 日爱因斯坦致 A. Sommerfeld 的信，见《爱因斯坦著作集》，第 8 卷，第 152 页。

[26] 派斯，《上帝是微妙的》，第 253 页。

[27] 1915 年 1 月 16 日爱因斯坦致 H. Lorentz 的信，见《爱因斯坦著作集》，第 8 卷，第 179 页。

[28] 引自派斯，《上帝是微妙的》，第 253 页。

[29] 爱因斯坦，"水星进动的广义相对论解释"（*Explanation of the Perihelion Motion of Mercury from the General Theory of Relativity*），1915 年 11 月 18 日，见《爱因斯坦著作集》，第 6 卷，第 113 页。

[30] 同上，第 117 页。

[31] 数学家 David Hilbert 独立地提出了另一个与此类似的方程。

[32] 引自 Clark，《爱因斯坦》（*Einstein*），第 200 页。

[33] 1916 年 1 月 17 日爱因斯坦致 H. Lorentz 的信，见《爱因斯坦著作集》，第 8 卷，第 179 页。

[34] 第二年，在一篇题为"广义相对论的宇宙学思考"（*Cosmological Considerations in the General Theory of Relativity*）的论文中，爱因斯坦对他的基本场方程进行了改进。他发现基本场方程似乎预示出宇宙是在膨胀的，所以他从方程（$G_{\mu\nu}$）左边的项中减去了另一个张量 $g_{\mu\nu}$，再乘以常数 λ。爱因斯坦坦陈，常数 λ 的值"目前还不知道"。这一结果能够保证广义协变性，以及爱因斯坦明确假定的宇宙有限性。这样一来，爱因斯坦的场方程 $G_{\mu\nu}=$

$-\kappa(T_{\mu\nu}-1/2g_{\mu\nu}T)$就变成了 $G_{\mu\nu}=-\lambda g_{\mu\nu}=-\kappa(T_{\mu\nu}-1/2g_{\mu\nu}T)$。爱因斯坦通过引入这一参数（纯粹是一个修正因子，也就是现在人们熟知的宇宙学常数），就能解释他对自己理论的预测，即宇宙是膨胀的。随后几年之内的时间里，爱因斯坦开始考虑这一概念的必要性，并于 1931 年从理论中永远去掉了常数 λ，并认为自己之前敷衍了事的做法是自己一生中“犯下的最大的错误”。70 年之后，为了解释超新星的测量数据，天文学家又重新恢复了这一参数。

[35] 西伊的文章，“爱因斯坦是骗子”（Einstein a Trickster）在 Jeffrey Crelinsten 的《爱因斯坦陪审团：相对论验证的竞赛》（*Einstein's Jury: The Race to Test Relativity*，普林斯顿：普林斯顿大学出版社，2006 年）一书中被引用，第 222 页。

[36] 有关日食的经典文章是 J. Earman 和 Clark Glymour 的“相对论与日食：1919 年英国日食探险和之前的探险活动”（Relativity and Eclipses: The British Eclipse Expeditions of 1919 and Their Predecessors），《物理学史研究》（*Historical Studies in the Physical Sciences* 11:1，1980），第 49~85 页。

[37] Alistair Sponsel，“构建‘科学上的革命’：推动人们认可 1919 年日食实验的宣传活动”（*Constructing a 'Revolution in Science': the Campaign to Promote a Favourable Reception for the 1919 Solar Eclipse Experiments*），《英格兰科学史杂志》（*British Journal For the History of Science* 35，2002），第 439~468 页。

[38] 1919 年 9 月 27 日阿尔伯特・爱因斯坦致保利娜・爱因斯坦的信，见《爱因斯坦著作集》，第 9 卷，第 98 页。

[39] 《自然科学》（*Naturwissenschaften* 7，1919），第 776 页。

[40] 引自 Clark，《爱因斯坦》，第 230 页。

[41] “皇家学会和皇家天文学会联合日食会议”（Joint Eclipse Meeting of the Royal Society and the Royal Astronomical Society），《天文台》（*The Observatory* 42，1919 年 11 月），第 389 页。

[42] 爱丁顿，《相对论》（*Relativity*），第八届哈尔登年度演讲，1937 年 5 月 26 日。

[43] 爱因斯坦，《想法和观点》（*Ideas And Opinions*，纽约：Bonanza Books，1954 年），第 311 页。

9

量子论的基本方程
薛定谔方程

$$\frac{d^2U}{dr^2}+\frac{2(a+1)}{r}\frac{dU}{dr}+\frac{2m}{K^2}(E+\frac{e^2}{r})U=0$$

说明：系统的量子态——例如，可解释为在特定位置探测到粒子的可能性——随时间而变化。
发现者：埃尔文·薛定谔（Erwin Schrödinger）。
发现时间：1926 年。

薛定谔方程是量子论的基本方程。该方程的研究在现代物理学中发挥了极其重要的作用。从数学的观点来看，薛定谔方程和数学本身一样，是取之不尽的。

——F. A. Berezin 和 M.A. Shubin，
《薛定谔方程》（*The Schrödinger Equation*）

从普朗克引入量子到薛定谔确认量子的普遍存在，科学共同体仅用了 25 年的时间。

1900 年普朗克首次提出量子的概念时，它不过是地平线上的一个小点。有了量子，普朗克就可以用经典理论解释黑体辐射。只要假定所有物体在吸收和辐射光的时候都是有选择性的（普朗克将这样的物体看做

是“谐振子”）——即按着一定大小能量的整数倍吸收和辐射能量，那么理论就仍然是适用的。许多科学家对此不以为然，认为这纯粹是胡编乱造，是在回避问题，而非真正意义上的科学。他们认为量子思想最终将被抛弃，并逐渐淡出人们的视野。

量子的持续拓展

不过在 1905 年的一篇关于光电效应的论文中，爱因斯坦拓展了这一思想。他提出，量子并不是基于谐振子的选择性，而是基于光的“粒子性”这一事实。在 20 世纪最初 10 年的末期，量子已经在物理学的各个不同分支中出现了。很多此前没把量子力学放在心上的人都开始注意到它。

1911 年，瓦尔特·能斯特（Walther Nernst）迈出了里程碑式的一步。能斯特是一位普鲁士物理化学家，起初他也像其他人一样对量子理论很不屑，认为它是“怪异”公式的产物。但是后来能斯特却利用该理论解决了汤姆逊所谓的“第二朵乌云”的问题，亦即将热的经典分子理论应用于涉及低温固体、气体和金属的实验结果上。能斯特宣称，经过普朗克和爱因斯坦之手（其实还应该提到能斯特本人），这一理论已经“卓有成效”，现在“科学界应该担起责任，严肃看待这一理论，并进行仔细研究”。[1]随后，在比利时实业家欧内斯特·索尔维（Ernest Solvay）的支持下，能斯特在布鲁塞尔组织了一次会议，引领科学家们从事这方面的工作。

这次会议是一个划时代的事件，它标志着量子（光以及所有其他形式能量的基本粒子性这一思想）将会成为科学的一部分。

与其他重大事件一样，这次会议的重要性也很快就显现出来了。与会者向那些未能参会的人们述说激动之情。诺贝尔奖获得者卢瑟福在回到英国剑桥后，向实验室的新成员——27 岁的丹麦人尼尔斯·玻尔“生动”地描述了讨论的情景。在巴黎，庞加莱写道：量子假设似乎引爆了

“自牛顿时代以来自然哲学史上最伟大、最彻底的一次革命”。[2]很多未能与会的科学家则通过会议论文集捕捉到了量子的灵魂。巴黎大学文理学院一位名叫路易斯·德布罗意（Louis de Broglie）的学生就是其中之一。德布罗意原本打算进入政府行政部门工作，不久前刚刚开始研究物理。他后来写道：论文集使他坚信要把自己的“毕生精力”投入到量子论中。

不过，量子却很难与牛顿力学相容，尽管它解决了很多关键问题。量子就像是一位无法被说服去参加活动的客人。而且这位客人即便是参加了活动，也会觉得尴尬，所以你必须要仔细关照他。我们不妨想象一下玻尔用量子去解释卢瑟福提出的到那时仍晦涩难懂的原子结构的思想。1911 年，卢瑟福提出原子就像是微小的太阳系，其中心的核（原子核）被电子包围着。不过，这与经典力学原理是相违背的：按照麦克斯韦的理论，绕着轨道运行的电子为什么不辐射出能量，落到原子核上呢？玻尔指出这其中的原因在于，按照量子的思想，电子只能以特定的量吸收和辐射能量，因此电子只能位于原子内部有限的静止轨道和能量状态上，并且只能吸收或者释放在这些状态之间跃迁所需的能量。这的的确确是个奇怪的假设。它表明原子中的电子［采用了美国哲学家威廉·詹姆士（William James）用来描述意识流的图像，这一图像可能影响了玻尔］在这些状态之间“停顿和发生跃迁”，并且在轨道之间跃迁时没有清晰的轨迹。[3]我们关心的是状态，不是轨迹——由此衍生出“量子跃迁”一词。玻尔将这一思想用于经典的原子测试用例——氢原子中。在氢原子中，一个电子绕着一个质子在轨道上运动。玻尔表明了他的这一假设如何能够预测出巴尔末公式。巴尔末公式是经验公式，它是由教师和数字命理学家巴尔末（Balmer）提出的，能够预测氢原子的谱线。[4]

这类既静止又运动的物体（现在只应用在光上，很快就会用到物质上）很快就产生了分类的问题。在经典物理中，即便是最小的物体也可分成两种类型：粒子和波。粒子是离散物体：每个粒子都具有特定的位

置和动量，并且在时间和空间上总是沿着一定的路径。而波是连续的：波从波源向外以球面的形式传播，没有特定的位置和方向，在时间和空间上慢慢展宽、变细。科学家采用不同的理论对粒子和波进行描述。粒子用牛顿理论描述。这一理论假定质量集中在一定的点上，受到力的作用，在任意时刻都具有特定的动量和位置。波用麦克斯韦理论描述。这一理论采用连续函数描述过程在时间和空间上是如何平稳演化的。这两种理论都得到了很好的发展，并且都具有确定性：只要输入初始状态信息，加以计算，就能预测出未来的行为。

那么，这类既静止又运动的物体应该归为哪一类呢？他们似乎既具有粒子的一部分特点，也具有波的一部分特点。这怎么可能？

1905 年，爱因斯坦在他的光电效应论文中给出了答案。他说：传统光学认为光是波，原因在于它所研究的光是大量的、是在时间上平均了的。但是，当光与物质相互作用时，也就是当光被发射和吸收时，在很短的时间尺度范围内，光就是粒子，被局限在一定空间，其能量是$h\nu$的整数倍（光“量子”，后来被称为“光子”）。爱因斯坦在给朋友的一封信中自豪地写道：这一思想是“非常革命性的”。[5]

接下来的 20 年间，物理学家们要么支持粒子论，要么支持波动论，并试图把各自的理论延伸，从而覆盖所有量子现象。

在粒子论一边，爱因斯坦扛起了理论的大旗。其实，他心里是并不怎么乐意的。在 1916 年的那篇重要的论文中，他拓展了光是以物理上实际存在的量子的形式吸收和辐射光的思想，并且提出每个量子都具有特定的方向和动量（$h\nu/c$ 的整数倍）。而且，爱因斯坦还提出了一个总的（可能有点夸大）的结论——他声称“球面波形式的辐射并不存在”。[6]现在，他已能说明在这一过程中能量是守恒的，因为从一端释放出的能量等于另一端吸收的能量。但同时，爱因斯坦也发现，要使理论成立，他必须以“概率系数”的形式引入统计学定律，来描述量子的发射和吸收。[7]他把这看成是一个令人不快的妥协，希望这种妥协只是暂时性的无奈之

举，并期望自己的工作不久就能为更深层次上的理解所替代。爱因斯坦在实验上的合作者中有康普顿。康普顿曾于 1923 年论证了“康普顿效应”，即电子与光子相撞后，二者反弹的方向都是固定的。[8]

波动论的胜利者之一是物理学家查尔斯·达尔文（Charles W. Darwin）。达尔文的爷爷是著名的博物学家达尔文——孙子沾了爷爷的光，不过他也对自己的这种角色感到有些不满。达尔文坚信光是以波的形式发射出来的。不过，他也知道，如果硬要用波动论去解释光电效应之类的量子现象，那么波动论将不堪重负。1919 年，达尔文写了一篇评论——“论物理学的基础”（Critique of the Foundations of Physics）。在这篇评论中，他预见未来物理学领域将会发生一些根本性的变化。他预言：量子现象可能会迫使物理学家放弃长久以来视为珍宝的物理学原理。达尔文半开玩笑似地写道：科学家们得接纳几近疯狂的想法，比如“赋予电子自由的意志”。[9]他认为：最不疯狂的做法就是保全波动论，抛弃单一事件的能量守恒，使能量守恒只在平均的意义下成立。

达尔文成为了玻尔的一位支持者。1924 年，玻尔联合了另外两人——亨德里克·克喇默斯（Hendrik Kramers）和约翰·斯雷特（John Slater）——要消灭掉爱因斯坦的激进思想，提出一种更加传统的方法，采用波动理论解释光的发射和吸收、光电效应以及康普顿效应。[10]这几位作者发现，要“谋杀掉”爱因斯坦的思想，得付出巨大的代价，的确，他们必须得抛弃能量守恒定律和通过可直观理解的方式了解光的发射和吸收机制的任何幻想。只有在平均意义下，能量守恒才是成立的。

“可直观理解”（visualizable，在德语里是 anschaulich）一词正是在此时成为物理学中的一个技术词汇的。理论中的东西要变成直观想象的或者通过直觉就能理解的，需要满足两个前提条件：理论中的变量要与实际事物联系起来，如质量、位置和能量等；理论的操作要与人们熟悉的操作联系起来，如逐点运动和超距作用等。因此，物体要可直观理解，并非一定要是牛顿式的。只要物体是在时间和空间上呈现的，那么即便

该物体是奇异的、非牛顿的，也仍旧是可以直观理解的。某个物体可直观理解仅仅意味着可以用像连环画一样的方式对其进行描述。连环画上的页面就相当于时间段，能确定在各个时间点的事件中的物体所处的位置。如果快速翻动书页，那么当前页上的内容就会与下一页上的内容平滑地融合到一起。

但是，认为玻尔–克喇默斯–斯雷特理论过于极端的（抛弃能量守恒和可直观理解），不光是大多数的物理学家，还包括这三位作者中的一位。斯雷特后来称自己签名是被迫的。在论文发表不到一年，玻尔–克喇默斯–斯雷特设想被实验否定之时，基本上没什么人感到意外。

玻尔–克喇默斯–斯雷特论文在科学史上是一份独一无二的文档。这篇论文很明显是错误的，但影响却很大，在史学家中声望颇高。它影响大的原因在于使粒子论和波动论之间的矛盾达到了顶峰。论文中提到：为了保存现有的一些观念，必须要作出这样的牺牲。各方的支持者此时只剩下谨慎和保守，尽力保全他们所认为的经典理论中最坚挺的元素。然而，来自量子现象的对抗却一刻也没有停止过。

的确，在量子理论出现后的最先二十五年的末期，它还是一团糟。历史学家马克斯·詹摩尔（Max Jammer）称量子理论是“假设、原理、定理和计算方法的一个可悲的大杂烩，而不是逻辑上连贯的理论”。所有问题的解答都与经典情形相仿，然后再通过一个加上了量子条件的“神秘的筛子”，筛除禁止的状态，留下几个允许的状态。这一过程涉及的并非是系统性的推导，而是“技巧性的猜想和直觉”，颇像“特殊的技艺或者甚至是艺术手法”。[11]从起点出发，要得出正确的状态，需要有理论的支持。也就是说，量子理论更像是一组指南，教你如何想出从点 A 到点 B 的路线。但实际上你真正需要的却是一张地图。

之后的 1925 年，两个观点截然相反的人各自作出了重大的突破，他们是海森堡和薛定谔。他们都努力想要保持理论的传统性，尽可能地保全经典物理框架。结果，他们却作出了革命性的突破。

时年 24 岁的海森堡哪怕是按照物理界的标准来看也是很年轻的。为拯救经典力学，他在自然的最底层就抛弃了它。他称：在原子中，不仅粒子和电子的轨道没有意义，甚至连位置、动量、速度、空间和时间等经典性质也都是没有意义的。而人们的想象力需要一个时空容器，如此一来都无法构想原子世界了。海森堡说：我们必须要把理论建立在所谓的“量子物理量”的基础之上，而这些量是无法可视化的。在第 10 章，将概述海森堡在提出他的这一方法时所经历的步骤。一次，海森堡注意到一个奇怪的问题：有几组量子理论量在它们所服从的特定“相乘”定义之下是不能对换的。也就是说，二者相乘的顺序将影响二者的乘积。起初，海森堡认为这个问题很棘手，想忽略掉它——不过他很快就发现，这正是量子力学的关键。1925 年，海森堡写出了“论动力学和力学关系的量子力学的重新解释”（On the Quantum-Mechanical Reinter- pretation of Kinematic and Mechanical Relations），为既没有粒子又没有波的量子计算提供了一种方法。该方法采用了所谓的矩阵数学方法，提供了一种正式的数学工具。只要往公式中输入数据，进行运算后，就能得到允许的状态。海森堡的导师波恩很快就发现了自己的学生重新发现了矩阵。不过正如其名，矩阵力学难以使用；而且很多物理学家对这样一个告诉他们“自然底层的结论无法预测”的理论也是有所抵触的。

时年 38 岁的薛定谔在物理学界算是年老的了。他的方式与海森堡的大同小异，不同之处在于薛定谔采用的是人们熟悉的经典力学工具：由他本人创建的波动方程，以及描述在时间和空间上平稳进行的事件的连续方程。在薛定谔看来，自然底层的事物是由完全可视化



埃尔文 • 薛定谔（1867—1961 年）

的波组成的。

走近薛定谔

1921 年，埃尔文 • 薛定谔来到了苏黎世大学（University of Zürich）。[12] 按照惯例，新来的教授要作一场公开的正式报告。薛定谔报告的题目是：“什么是自然法则？”在报告中，他称“自然界的规律很可能都具有统计特征”。[13] 的确，为描述诸如气体等由大量微小物体组成的系统的行为时，麦克斯韦把统计学定律引入了物理学中。这些定律虽然方便，不过他们仅仅是一种近似，并不精确。而且另一方面，人们的知识是有限的。原则上说，预测整个系统的性质，需要对每个分子的运动特性一一进行跟踪，把力和质量的值输入牛顿定律中，加以计算，就能预测出系统在过去和将来的行为。1961 年，爱因斯坦第一次在文章中用到概率时，认为这种做法只是暂时的。薛定谔不知道，自己将要做的很快就会被解读为把统计学永远引入到了自然定律中（自然定律之下不再存在更深层的基础定律）。

因为疾病的原因，薛定谔的工作推迟了几年。不过 1925 年之前他一直都在进行量子论方面的研究，同时也参加了由本校和附近的瑞士联邦理工学院联合举办的学术讨论会。1925 年秋的一天，ETH 方面的组织者之一——荷兰物理学家彼得 • 德拜（Pieter Debye）邀请薛定谔就路易斯 • 德布罗意新近发表的论文作一个报告。德布罗意是一位刚刚毕业的法国物理学家，他在阅读了索尔维会议记录后决心投身到量子论的研究中。他提出了电子伴随着波动过程的概念，利用普朗克定律 $E=h\nu$ 把电子的动量与波长结合了起来。以该假设为基础，德布罗意就能解释旧量子论的量子化条件。因此，在下届的 ETH 学术讨论会上，薛定谔将负责介绍这位年轻法国物理学家的思想：如果假定电子具有整数波长，就能得到正确的轨道。

坐在前排的德拜却不以为然，觉得上述思想“相当幼稚”。按照惯例，能坐在前排的都是名人。德拜说：如果某个物体是波，它就要有一个恰当的波动方程。

德拜的意思似乎是：波一般指的是正在发生波动的物体。实际上，在物理学的其他领域，波是“某物体”的运动方程的解。德布罗意指出了与电子相关联的波，但是不能说明究竟是什么在波动，也不能给出波动方程。

与参加 ETH 学术讨论会的大多数人不同，薛定谔认真考虑了德拜的评论。同时他也受到了爱因斯坦对德布罗意工作的评述的影响，即“所有运动都与波动场相联系”。[14]在完成了正在写的气体量子论的论文后，薛定谔就和一位旧时的女友一起去阿罗萨滑雪度假了。这位神秘女郎的身份，一直以来都是一个谜。他 1925 年所写的日记已经丢失，几个明显的可疑人物也都被一一排除。薛定谔的同事曾这样评论他：“薛定谔的伟大工作是在他生命中一段情欲旺盛的时期做出的。”——这个评论令人费解。不过 38 岁的年纪对于搞物理的人来说虽然不算小了，可要说到情欲，却绝对不算什么。薛定谔的传记作家这样写道：“就像激发出莎士比亚产生灵感，写出十四行诗的那位黑衣女士，阿罗萨的这位女士永远都是一个谜。不管激发出薛定谔灵感的到底是谁，事实是薛定谔的精力从此大增，开始了长达 21 个月的持续创造性活动。直至今日，在整个科学史上也无人匹敌。”[15]

12 月 27 日，薛定谔从阿罗萨写了一封信给维恩：

> 现在我正在努力研究一种新的原子理论。要是懂的数学知识再多点就好了！不过我很乐观：只要能解出来，这个理论就一定是非常漂亮的。[16]

1 月 9 日，从阿罗萨回到苏黎世的薛定谔，似乎还在攻克这一课题。不过不久后，他举办了另一场学术讨论会，并说：“我的同事德拜建议说

应该要有波动方程，我现在已经找到！”1926 年，薛定谔发表了一系列总数 6 篇的非同凡响的论文，把波动方程呈现了出来，还讨论了方程的意义。这 6 篇论文“用特征方程问题处理量子化”分 4 部分发表，“无疑是科学史上最具影响的贡献之一”[17]。其中有一篇是关于量子世界和经典世界之间的过渡，另一篇是关于波动力学和矩阵力学之间的关系——薛定谔展示了他的波动方程，并指出了方程的意义。

薛定谔在薛定谔方程中引入了波函数，并把该函数称为“新的未知量 ψ”。与德布罗意一样，他分别将波长与动量、频率与能量联系在一起。他提出：原子世界的行为是由 ψ 场的波组成的。薛定谔最初感觉 ψ 场就好比是电荷密度、一种粒子云。这些波可以相互叠加、干涉、产生节点等。他在共计四部分的文章的第一部分中指出，这些“明显可见”的图像使人们可把实验中观察到的现象描述为“两个原子或分子相撞后如何反弹，或者电子或 α 粒子在射穿原子后如何改变方向”。[18]他说，实际上由此就可以把原子中电子的状态视为驻波。驻波就像是小提琴的弦，在振动时保持基本形状不变。他在文章的第二部分提到，希望自己能够说明，他的理论可以表明波群或“波包”如何在“所有方向上都以较小尺度”形成，“同时遵守与机械系统中单个像点一样的运动定律”，即与单粒子的行为相同。[19]

不过事情并没这么简单。ψ 场毕竟是一个数学量，只有在补充了其他一些条件后，它的性质才会决定观测到的物理性质。薛定谔加入这些条件后，发现条件中包含复数 i。起初，他对波动理论中出现复数感到心烦意乱，并试图去掉它，但最终失败了。他非常苦恼，因为复数包含两部分——实部和虚部。复数的出现表明波函数有一个相位是无法直接观测到的。相位就像是钟表，它是不断循环往复的。相位中有虚部就表明有一部分相位是无法直接测量到的。它在时间上的振动无法从外部“现实”中观察到。薛定谔方程描述了多维空间和组态空间中的波动现象。

究竟什么是 ψ？薛定谔在第一篇文章中写道：起初他希望“将函数 ψ

与原子中的一些波动过程联系起来。这是一种比电子轨道更贴近现实的方式。如今，电子轨道是否存在仍有很多疑问”。也就是说，这一波动过程会产生与物理学家所谓的量子态相似的东西。量子态是离散的、不连续的，也不是由连续过程构成的。在第一篇文章的末尾，薛定谔说：“想象一下，在量子跃迁过程中，能量从一种形式的振动变化成为另一种形式。这种设想比电子跳跃要令人满意得多，因为波动形式可以在空间和时间上连续地发生变化。”同时，他还将自己的结果与玻尔、克喇默斯和斯雷特失败的波动说进行了比较。不过，薛定谔说自己现在不会去追随这些思想，也不会为了让自己满意而采用“中立的数学形式”[20]来呈现思想。

进一步，他希望波的叠加能够产生一种固定不变的“波包”，就像水池中具有稳定清晰波峰的行波一样。这样就能解释 ψ 场在表现为单粒子时所发生的现象。

不久，薛定谔就发现这种直接的“波包”感觉是不可能的。在“量子化”论文系列的第四部分，也是最后一部分中，他写道：“ψ 函数本身不能，也不可以直接用三维空间解释。只有在奇异正规空间或者“构形”空间中，波函数才是波。不过，薛定谔的方法还是达成了他的目标：采用物理学家熟悉的、容易上手的方程，从宏观世界讨论原子世界——时间、空间、波等。薛定谔的原子世界图像基本上还是直观的。一个很重要的方面是如何用人们对世界的常识概念来解释这些方程。如果还能防止那种坚持认为是上帝或者其他超自然的事物在推动粒子的运动，而放弃理性计算的观点，就更好了。

但是其他人很快就扭曲了薛定谔的直观解释。

波的解释

1926 年夏天，哥廷根物理学家波恩发表了他在原子碰撞（类似于电子与原子的碰撞）方面的论文。波恩是海森堡的导师，也是矩阵力学的

奠基人。碰撞在经典物理学中处于核心地位。它也被波恩认为是理解原子领域的关键问题之一。他曾试着用矩阵方法来做，但没有成功。不过波恩却得出一个令人吃惊的结论。他宣称：“为此，只有薛定谔的形式主义可以证明它自身的正确性。基于这一原因，我倾向认为它是量子论中最重要的公式。”[21]不过波恩也有一些海森堡不愿听到的消息。他完全不能理解海森堡的断言——ψ 函数指的是电子的电荷密度。波恩认为：薛定谔方程并没有提供事件状态的信息，而是给出了事件状态出现的概率。薛定谔方程所描述的可在空间中连续运动，并与势相互作用、产生干涉的 ψ 函数并不是实际的场，而是概率。“我们把力从直接决定粒子运动的经典角色中解放了出来，让它们可以决定状态的概率。”[22]波恩很奇特地将波动力学和矩阵力学结合到了一起：它一方面体现了连续性和因果性，另一方面也体现出了不连续性和概率性。“粒子的运动服从概率定律，而概率是按照因果律传递的。”[23]

几个月后，波恩以前的助理沃尔夫冈·泡利（Wolfgang Pauli）给出了另一种解释。与波恩把 ψ 解释成状态的概率不同的是，泡利认为它是粒子的概率——ψ^2 表示电子在特定位置上出现的概率。这离薛定谔对函数的解释就更远了，因为 ψ 函数的实际意义完全不存在了。ψ 函数指的是某物发生的概率，不是事实——比如计数器的咔哒声或者粒子的存在性等。要了解事实，就要搭建起设备，将亚原子粒子的性质表现出来。这是世界与波函数的描述之间的相互作用。

波恩–泡利解释把粒子论和波动论相结合，很快就为大多数物理学家所接受。但是这种结合也有代价。这个奇怪的代价就是这两种理论的特征都消失了。如果用牛顿定律研究粒子，粒子就是可观察的，定律也是确定的——从粒子的初始状态出发，通过定律加以计算，就能预测出粒子未来的行为。麦克斯韦定律在应用于波时也是如此：波是完全可以观察的，波的性质也可一一测定。麦克斯韦定律是确定性的，能够描述波随时间的变化。也就是说，波动论和粒子论都是关于可预测量和可观测

量的。现在，波恩–泡利的解释把粒子论和波动论结合了起来，不过二者各自都有一些牺牲。薛定谔波在构形空间中波动。粒子虽然能观察，却不能预测；而波虽能预测，却又无法观察。如果观测到了物体的位置和动量，就无法预测物体下一步的状态。波是用来预测另一事件发生的可能性的，但在事件被观察到之后，波就没有什么价值了。此时，必须要丢弃波或者对波进行“重置”，修改并加进新的信息。

现在，对这一解释的呈现常常是误导性的。我们常能听到，测量完毕后说波函数“坍缩了”，而不是波函数被丢弃或者重置了。这个比喻抓住了一点：事件发生前（比如在探测到粒子之前），粒子可能在任何位置，所以人们认为事件或粒子是处处存在的。由此产生的图像是突然在某一点被吞噬的、在空间中延伸的结构。这一图像的确形象生动，但却具有迷惑性。其实，波只是一种概率，而不是一个“物体”（这也是所谓的导航波理论——“万物均不坍缩”的优点之一，波仅存在于已在波当中的粒子之上），波函数的作用只是给出概率。它是以可预测的、确定的方式不断流动的。不过一旦有事件发生，函数的作用就消失了，必须重置。

纽约石溪分校的物理学家阿尔弗雷德·戈德哈伯（Alfred S. Goldhaber）重新考虑了约翰·惠勒（John Wheeler）对爱因斯坦的广义相对论的著名评论，对薛定谔方程做出如下评论：“波告诉粒子往哪里去，粒子则告诉波从哪开始，到哪结束。”

颇具讽刺意味的是，这种原本打算设计成符合人们直觉的方法，却常常给出错误的图像。

薛定谔方程表明世界上的事件会发生剧烈的变化。最终，以下情况将不再成立：向方程中输入数字，通过计算方程，就能得出预测结果。与此相反，输入数字，并计算方程，得出的将是概率——得到了一定位置上事件发生的概率。最好的办法就是：尽可能少的信息和尽可能少的计算步骤。我们也不可能为事件建立一个相册。因为如果不停地让一件事件重演，粒子将会在纸上不同的点出现，其位置是这些点的平均值。

薛定谔本人从来没想到过这种解释，认为这种解释其实是一种“折中”[24]。他写道：这样解释倒是很“方便”，不过不会“这么轻易就到头了”。他坚持认为应该继续寻找解释，找出因果机制。他指出，一些外表合理的现象其实是“非常荒唐的”，这中间就包括现在非常有名的“薛定谔的猫”。此假想实验是：把猫锁在一个盒子里，盒子里有一个残忍装置。同时盒子里还有放射性物质。原子衰变后，将触发“恶魔般的装置”，将猫杀死。因为衰变是由 ψ 决定的，于是乎猫的死活就也就由 ψ 决定。这样似乎就可以得到这样的结论：猫的存在状态是叠加的，即半死半活。结论虽然不符合常识，不过这个聪明的想法却说明将微观世界的理论推广到宏观世界是有问题的。

然而，虽然该领域的所有工作者都在使用薛定谔的波动方法，他们却忽略了薛定谔关于现实中波的结构的前提。海森堡在哥廷根的合作者给出的解释成为人们最喜欢的版本。历史学家马拉·贝勒（Mara Beller）写道“薛定谔的方法是不可缺少的，不过基本原理并非如此”。[25]

波恩对薛定谔方程的解释改变了人们对世界完整理论的认识。一般来说，人们在具有了完整理论后，就能告诉自己一些现实的东西。物理学家所教授和使用的大部分理论都有投机取巧的成分，不符合人们的上述预期。这些理论并没有给出一个完整的图景。它们给出的只是任意一个现实理想化后的情形。例如，理想气体定律忽略了一些人们已经充分理解了的力，如范德华力和硬核排斥力。不过人们并不在意抛开现实情况，去讨论理想情况。原因是后者应用起来要方便得多。我的一位同事将此称为“无害的捏造理论”。这些理论之所以没有什么害处，是因为这些限制并不会威胁到世界的一般假设。我们对自己的投机取巧心理是清楚的，也知道这并不会影响人们对世界的认识。

然而，波恩对薛定谔方程却有着不同的解释。它使我们认识到人与世界相互作用的方式会影响人向世界中输入的内容。它使我们明确意识到进行测量就是在与世界进行交互。“人只是媒介”的思想开始清晰地进

入人们的视野——人们并不是观察舞台上发生的事情，然后把它记录下来，而是在参与舞台演出。薛定谔方程的主流解释将这一思想置于重要位置，这点是经典物理所做不到的。其实，决定究竟是牺牲还是进步的是人们自身的期望。量子力学的这一巨大飞跃确实要求从本质上重新思考世界：怎么才算是“理解”自然，以及如何描绘“现实”。而且，描绘的方式是海森堡所谓的，当时许多科学家都在经历的牺牲——这种牺牲是痛苦的，所以他们奋力挣扎，不想看到这种牺牲。直到现在，许多人依旧如故。



茶　歇

科学家的双重意识

> 一个使他在其中失去了自我意识的世界，在这个世界里，他观察自己时，完全依赖于其他被揭示的事物。这是一种特别的感觉，一种双重意识，一种总是用别人的视角看待自己的感觉，一种用面带讥笑的世界的录影带来衡量自身的感觉。
>
> ——W.E.B.杜波依斯（W.E.B.Dubois）

海森堡和其他一些矩阵力学支持者不仅怀疑薛定谔的工作，甚至还对其充满敌意。这些人就是贝勒所谓的“教条地认为旧的概念就是好的，而不去作冷静、客观的判断的人”。与笨拙、复杂的矩阵方法相比，薛定谔的方法更加优美、简洁。这也令他们越发地恼怒。[26]而薛定谔也毫不掩饰自己对矩阵力学的蔑视。的确，量子力学的形成是科学史上的一个标志性的篇章：双方都有敌对情绪，私下里剑拔弩张，就是在正式发表

的论文里也常带有“不寻常的情绪化倾向”。[27]从历史资料中不难发现嫉妒、竞争、愤怒、怀疑、信念、紧张、希望、绝望和沮丧的情绪。然而在历史记载中，科学研究里的这种情绪化倾向却常常被省略掉了。

大多数科学史都是这么写的：有人作出了一个意外发现；于是人们提出了各种解释，不过都不合适；然后又搭建新设备再次测量，但解释仍不完整；之后又从另一个角度，采用其他设备对现象进行测量，如此等等。

这就是可以称为解释科学如何发挥作用的“标准模型”。该模型强调的是科学的集体共有性和客观性，没有考虑特定个人的感受。它的基本组成就是发现、仪器设备、测量和理论。科学概念被认为就是揭开谜题、控制自然。科学家们对这一模型很熟悉，所以他们在介绍自己的工作时往往也是强调上述内容，从而进一步加强了标准模型。在该模型中，情感生活和情感体验以及个人的成功和失望等似乎是一辆行驶在某条轨道上的列车；而科学家的职业——研究项目、所作出的发现和事件等，是沿着另一条轨道行驶的列车。两列列车分别由不同的车头牵引，行驶在各自的轨道上，构成了科学家的两面。

不过，如果用心去聆听科学家讲述自己的工作（很多知名传记作家也是这么做的），也能听到另外一种强调个人感受的讲述。作出发现时的激动，解释无效时的彷徨，考虑如何解释时的好奇心，尝试诸多解释都无效后的困惑，设计新仪器设备时的想象力，“柳暗花明又一村”时的惊喜，甚至是作出根本性发现时的惊愕都可以成为激励因素。科研过程变得浑然一体，反衬出标准模型的局限性。也就是说，在标准模型之外还有其他东西，它注定最终要被取代。取代它的就是大统一的方案：在该方案中，两列列车合二为一，科学是由兼顾生活和工作的个人一手完成的，而不是分工完成。

为清楚起见，不妨看一下几年前费曼发表的书信集。你会发现费曼的性格、做事习惯等都是同他的职业紧密交织在一起的。他的好奇、推测、演讲和清楚地向人们讲解科学的愿望完美地结合在一起。费曼既是一名物

理学家，也是一名教育工作者，两种角色与他本人的性格合而为一。他对科学的洞察力和与他人的互动也都因此而收益。费曼写道："生活真正的乐趣在于不断接受检验，去发现自己有多大的潜力。"

从爱因斯坦身上也可以看到这种"大统一"。科学史学家杰拉尔德·霍尔顿（Gerald Holton）曾写过一篇很好的文章，叫作"爱因斯坦的第三天堂"（Einstein's Third Paradise）。他的"第一天堂"是幼年时所经历的"具有强烈宗教意识的阶段"。爱因斯坦把它称为"少年时的宗教天堂"阶段，他和姐姐玛佳就是两个例子。12 岁的时候，爱因斯坦的"第一天堂"阶段结束了。此时，他读到了大众科普书籍，通过这些书知道了圣经中的故事不全是对的。他在读了别人送的一本欧几里得平面几何的小册子后，从中找到了乐趣。爱因斯坦把这本小册子叫做"圣书"、"奇书"。他还读到了其他科学著作。通过这些著作，他知道世界是一个"伟大的、永恒的谜"。思考这个伟大的谜会给他带来"内心的自由和安全感"。他将此也称作是"天堂"。爱因斯坦这样写道：离开第一天堂，进入第二天堂是一次"将自己从'纯粹个人化'的锁链中，从满是愿望、希望和原始情感的状态中解脱出来"的尝试。

传记作家们倾向于把这两个天堂相互对比，认为它们是爱因斯坦一生中的两个彼此分离的阶段：从宗教期到非宗教期。但是霍尔顿不同意这种说法。他从爱因斯坦成熟的内心中，看到了两个天堂的融合："爱因斯坦一生所作出的杰出科学工作的意义所依赖的正是残存下来的年轻人的炽热虔诚。"

在第三天堂中，爱因斯坦似乎代表了这样一个人：他具有人们所谓的宗教式的情感，这种情感对他的工作至关重要，可他并不认为存在这样一位"万物之主"（Master Mechanic）。正如爱因斯坦在一封信中所写的，他"根本不相信宗教"。第三天堂是可以用我所说的大统一来描述的一类事物。回想一下爱因斯坦在纪念普朗克诞辰六十周年上的讲话。他说：找出世界的简单明晰的图像，不仅是科学的目标，还是人们深层心

理上的需要。科学家在致力于追寻这一目标时，会把它作为“情感生活的重心”。他还说，研究最棘手的科学问题，需要有“一种与宗教人士或者情人的感觉类似的状态”。霍尔顿曾提到爱因斯坦和其他一些人在科学发展过程中所经历的情绪上的起起落落。这些人的心理投入与他们为自己所设定的任务不能割裂来看。科学与个人投入是紧密相连的。爱因斯坦着手将看似不同的现象统一到一起的做法，被霍尔顿看作是情感生活与职业工作相互渗透的典型。在 1901 年写给格罗斯曼的信中，爱因斯坦提到了毛细作用（涉及他的第一篇论文），指出物体的对立行为如何联系在一起。他在给康德的回信中说：“在直接感官看来完全不同的各种事物，却能统一起来。认识到这点真是一种奇妙的感觉。”15 年后，在另一封信中，爱因斯坦说他“概括事物的动力是自己本身的需要”。霍尔顿指出：实际上“爱因斯坦也有统一事物的强烈冲动”。他厌恶民族主义，希望能建立一个全球统一的政府。霍尔顿总结说：“没有界限，没有障碍——不但生活中没有，而且现实世界中也没有。爱因斯坦的生活和工作如此完美地交织在一起，于是人们顺理成章地认为他在研究大统一的方案（相互融合，达到一致）时二者是齐头并进的。”同样，霍尔顿还说：“爱因斯坦的科学和宗教情感中没有界限和障碍。”晚年的爱因斯坦在写到科学和宗教时，常常用相同的措辞来表达科学和宗教的目的。“我认为宇宙宗教情感是科学研究最强大、最高贵的动力……一位与我同时代的人说：在如今的物质时代，严谨的科学工作者是仅有的极度虔诚的人。这种说法并不过分。”而且，“人们所能拥有的最美丽的体验就是神秘感。这些人们无法参透的关于事物存在的知识、对最深奥的推理和最光芒四射的美的感受，只有以其原始形式出现时才能为人们所理解。也正是这些知识和情感产生了虔诚。从这种意义上说，并且也只有从这种意义上说，我才是一个虔诚的教徒”。于是，从爱因斯坦身上，我们依旧能够瞥见标准模型之外的东西：科学中，性格和个人情感既不是科学过程的边缘，也不是科学工作的序幕。正是它们支撑起了科学，使科学不断向前发展。

注 释

[1] 能斯特，引自马克斯·詹摩尔，《量子力学概念的发展》（*The Conceptual Development of Quantum Mechanics*，纽约：麦格劳-希尔，1966年），第59页。

[2] 詹摩尔，《量子力学概念的发展》，第170页。

[3] 同上，第178页。

[4] 巴尔末在论文中预言："在我看来，氢原子注定要从其他物质中脱颖而出，开启人们了解物质结构和物质属性的大门。"引自詹摩尔，《量子力学概念的发展》，第65页。

[5] 爱因斯坦给 C. Habicht 的信，1905年5月，见《论文集》，第5卷，第20页。

[6] 爱因斯坦，"论辐射的量子论"（On the Quantum Theory of Radiation），见《论文集》，第6卷，第220～233页。

[7] 同上。

[8] 《物理评论》（*Physical Review*）21（1923），第483～502页。

[9] 詹摩尔，《量子力学概念的发展》，第171页。

[10] 玻尔，克喇默斯和斯雷特，"辐射的量子论"（The Quantum Theory of Radiation），《哲学杂志》（*Philosophical Magazine*）47（1924），第785页。

[11] 詹摩尔，《量子力学概念的发展》，第196页。

[12] 数位科学史学家广泛分析了波动方程非凡的发现过程，其中包括 Martin Klein 的"爱因斯坦和波粒二相性"（Einstein and the Wave-Particle Duality），刊于《自然哲学家》3（1964），第3～49页；L. Wessels，"薛定谔是如何发现波动力学的"（Schrödinger's Route to Wave Mechanics），刊于《哲学科学研究史》（*Stud Hist Phil Sci*）10（1979），第311～340页；詹摩尔，《量子力学概念的发展》，1966年，第5章，第3节。

[13] Walter Moore，《薛定谔的一生》（*The Life of Erwin Schrödinger*，剑桥：剑桥大学出版社，1994年），第195～196页。

[14] 引自 Mara Beller，《量子力学解释的起源：1925—1927》（*The Genesis of Interpretation of Quantum Mechanics 1925—1927*），博士论文，马里兰大学，1983年，第124页。

[15] Moore，《薛定谔的一生》，第195～196页。

[16] 出处同上。

[17] 詹摩尔，《量子力学概念的发展》，第267页。

[18] 薛定谔，《波动力学论文集》（*Collected Papers on Wave Mechanics*，普罗维登斯，罗得岛州：AMS/切尔西出版公司，1982年），第59页。

[19] 同上，第 20 页。

[20] 同上，第 9 页。

[21] 詹摩尔，《量子力学概念的发展》，第 284 页。

[22] 波恩，“量子力学的物理问题”（Physical Aspects of Quantum Mechanics），《自然》，119（1926），第 354～357 页。

[23] 引自詹摩尔，《量子力学概念的发展》，第 285 页。

[24] 引自 Beller，《量子力学解释的起源：1925—1927》，第 144 页。

[25] Beller，《量子力学解释的起源：1925—1927》，第 105 页。

[26] 引自 Mara Beller，《量子力学解释的起源：1925—1927》，第 86 页。

[27] 同上，第 91 页。

10

与不确定性同行
海森堡不确定性原理

$$\Delta x \Delta p \geqslant \frac{h}{2}$$

说明：要准确描述小空间中粒子的位置，粒子的动量就会变得不确定，反之亦然。粒子总体的不确定性大于等于某一确定量。

发现者：海森堡。

发现时间：1927 年。

人人都理解，或者认为自己理解不确定性。

——迈克·弗雷恩所写剧本《哥本哈根》中的海森堡台词

说起来我们亏欠海森堡的有很多。作为量子力学的奠基人之一，他给物理学留下了一笔巨大的遗产。而作为不确定性原理的提出者，他又给物理学之外的其他领域留下了同样巨大的遗产。爱因斯坦或许受到公众更广泛的认同，他的相对论是在流行文化中一下子出现的。然而对公共话语和流行文化，海森堡有着与爱因斯坦同样深远的影响。

大部分非科学人士都认可爱因斯坦的方程 $E=mc^2$。一般而言，他们都清楚该方程的效应只在某些约束条件下才会显现，只有物理学家才真正清楚它的含义。而海森堡方程 $\Delta x\Delta y \geqslant h/2$ 却不是这样。海森堡不确定性原理对大众而言似乎具有精神上的意义。这种意义很快就变得深刻和明显起来。比如，在任何一个书店的近现代区，都可以见到自信地坚持不确定性原理的断言：不确定性原理的含义是迷幻的，预示着一场“文化上的革命”。即便在学术圈中，对这一原理的各种奇怪解释也频频露面。下面是一段发表在《美国戏剧》（*American Theater*）上的著名戏剧导演安·博格（Anne Bogart）和著名声乐教练克莉丝汀·林科雷特（Kristin Linklater）之间的对话：[1]

> **林科雷特**：有些思想家说最伟大的精神境界是不安全感。
> **博格**：海森堡已经证明了这一点，从数学上。
> **林科雷特**：没错。

但是我们到底在哪里？是怎么到这里的？不确定性原理起源于一种纯数学的原子物理研究方式。在原子物理领域，不确定性原理有着明确的定义，而且适用范围也有着严格的限制。

通向赫尔戈兰之路

海森堡的父亲是德国慕尼黑大学的一位希腊语教授。海森堡本人具有诗人般的个性：他长相时髦潇洒；但身体虚弱，很容易过敏；具有优秀的音乐才能；对周围世界的反应常常是敏感的，带有情绪化的色彩。[2]他颇有想象力、智力超凡，喜欢用非常规的严格数学方法使理论与实验相吻合。海森堡也非常幸运，他身处一个非常热烈、非常渴求的科学团体中，并从中汲取营养。该团体的成员有玻尔、波恩、约尔当、克喇默斯和泡利。

这些理论物理学家大体上分布在三个研究中心：慕尼黑、哥廷根和哥本哈根。每个中心都有各自的特色。慕尼黑主要以实验为主，哥廷根是世界闻名的正规数学中心，而哥本哈根有着研究量子世界的强大哲学方法（源自创始人和领导人玻尔）。这一物理学家团体中的激烈，甚至常常是异常坦率的交流（私人谈话、工作进展的信件和手稿，以及发表论文的复本）保证了所有参与者都有很高的水准。某个人提出的想法常常是由另外一个人实现的。作为中心人物，海森堡经常在三个中心之间穿梭。他的深刻见解也时常在谈话中显现出来。



海森堡（1901—1976 年）

1923 年 7 月，海森堡在慕尼黑完成了博士考试，并计划于当年秋天去往哥廷根，在波恩的指导下工作。谁曾想，被公认为神童的海森堡竟然几乎完全忽视了实验物理，差点没能通过考试。他甚至连蓄电池的工作原理都不能解释。后来，多亏一位主考官的强行干预，海森堡才勉强通过。丢脸的考试之后的第二天，带着一丝沮丧的海森堡突然在哥廷根出现在波恩的门前。他向波恩承认了自己考试的窘况，问波恩是否还愿意把自己留下。波恩的答复是继续支持他。于是海森堡放心地离开了，

去参加每年必去的一个青年组的夏日旅行。

此时正是历史学家马克斯·詹摩尔所描述的量子理论乱作一团的时期。如今，问题（最终被完全解决）所采用的分析方法还是传统方法，加上量子条件，得到几个“允许的”运动状态。当时，年仅 21 岁的海森堡决心要把这一切合理化。

海森堡知道一定要从经典力学出发。他后来说道：“量子力学的概念只能用已知的牛顿概念加以解释。也就是说，量子力学是以经典物理学为基础的。这是玻尔极力强调的一点——我们不能脱离已有的经典物理学去谈论量子物理。”[3]

在经典物理学中，所有事件都是在四维时空场中发生的。在任意时间，所有物体在空间中都有一个特定的位置。物体之所以从一个地方运动到另一个地方，是因为它对一定的力作出了响应，并采取了一定的路径。经典物理学主要关心物体受到扰动之后会发生什么现象，追踪何种力会产生何种效应。与平稳流动的水流一样，所有物体运动的路径都可以被追踪和预测，物体将从每个点连续平稳地运动到下一个点；而物理性质（都可以测量）在空间机械地在时空场中连续平稳地传播。因此，经典物理学为宇宙的终极元素组成和这些元素之间的相互作用提供了一种自信的实体论（或图景）。这种事件就是图景（anschaulich）。

然而，25 年来人们不断提出量子现象经典模型的尝试却全部失败。受所在丰富智力环境中各种辩论的激发，年轻的海森堡开始怀疑问题可能并不是出在这上面。如果构建原子世界图景（电子的位置和路径及其轨道的大小和频率）的尝试从一开始就注定要失败呢？他听到泡利曾说过，原子事件模型只具有“象征意义”，是量子现象的经典“类似物”。[4]麦克斯韦得出麦克斯韦方程的过程不就是一个教训吗？有时候，只有抛弃机械解释，才能捕捉到真实。难道没有这种可能：理论家基于实验家测量结果构建模型时，模型仅仅是人们无法描述的现实的象征？[5]海森堡曾写道：“科学上的进步常常会伴随着牺牲。这种牺牲的代价就是我们

声称要理解自然。”海森堡和同事们想，这次可能会是真正的“直观”了。

于是海森堡决定继续下去，他不再尝试去提出描述原子级事件在时空中如何发生的理论。利用在哥廷根学到的对形式结构的鉴赏能力，他要为实验学家实际观察到的现象找到纯数学的描述：电子所发射出的光的频率和振幅。这些描述需要遵守对应原理（大的量子数遵守经典定律）和其他一些约束条件，如能量守恒原理等。不过，可测属性和连续传播函数是不需要的。实际上，在海森堡看来，不连续性是量子世界的一个主要的独特特征。

这一发现非常重要，就好像是哥伦布发现了太阳系的结构。两者都改变了科学家惯常看待世界的视角。他们常常会把自己天真地以为是客观实在的反映的事物当成是人类观察者和自然之间相互作用的复杂产物。

这一步也是革命性的，不过通向这一步的道路却并非全由海森堡一人建成。第一，他利用了从玻尔、波恩和其他人那里获得的理论工具，并且抛弃了时空场。抛弃的理由仅仅是为了能使这些理论工具派上用场。第二，海森堡所采用的策略已有一个很好的先例——爱因斯坦在 1905 年时就已采用这一策略提出了狭义相对论。爱因斯坦抛弃了把传统意义上的“同时”当成是“在统一的时空瞬间发生”的做法，转而采用观察者能够观察到的形式对其重新定义。海森堡希望，通过抛弃原子内部的“位置”和“动量”等传统概念（不能观察到，是推断出的量），用实验者从外部观察到的现象（谱线的频率和振幅）来对其重新定义，从而取得同样重大的突破。最终，所有的尝试却都归于失败。虽然构建理论的尝试已被放弃，但这一激进的做法仍不足以描绘出观察不到的那些事物。

然而，与大多数革命一样，这一做法也有着长期的影响。这种影响需要数年时间才能变得清晰起来。如果说成为一个“事物”意味着在特定时间占据特定的位置，那么这种做法就意味着“从原子领域消除粒子的概念，即‘物性’”。[6]它实际上是替代了牛顿的自然实体论。在自然

实体论中，物体最基本的部分在特定的时间都是客观存在于特定位置上的。而新的实体论，如后来一位科学哲学家所说的，涉及的是“科学工作核心层面上的微妙的主观性”。[7]主观性与如下事实有关——我们对原子世界的描绘并不是客观实体的反映，它在一定程度上是构建图像的人类思维中的函数。这种微妙性与“思维所起的作用目前尚不清楚”这一事实有关。

当时不清楚的东西有很多。1925 年的头几个月，海森堡的思路开始时断时续地铺展开来。他和哥廷根的克喇默斯利用不包含经典变量，只含有频率和振幅的方程合著了一篇论文。后者的贡献是一个重要的线索。克喇默斯指出，只有这些频率和振幅与状态对联系起来的时候，才能得出正确的矩阵。后来，海森堡在滑冰时受了伤，在慕尼黑疗养了几个月的时间。他去了哥本哈根和哥廷根，还去了山上旅行，4 月底才回到波恩的研究所，准备教暑期班。之后，海森堡已经准备好用纯数学的形式改写玻尔对电子动量（p）和位置（q）的描述。不过他并没有把计划告诉导师，而是像玻尔所说，“对外保密”。[8]

后来，海森堡有一次回忆说：“我沿着这些思路所做的工作不但没有受到那次不幸挫折的影响而延迟，而且还向前推进了。”[9]5 月，海森堡出现了严重的花粉过敏，他向波恩请了两天假。波恩同意了。于是，海森堡就前往赫尔戈兰。赫尔戈兰是北海上的一座岩石孤岛，不适合草和其他产生过敏原的植物的生长。他离开的前一天晚上，指给他房间的旅馆女店主看到他浮肿的脸时大吃一惊，以为海森堡被人打了。在岛上的时候，一旦工作条件具备，他就开始试着验证自己的想法是否符合能量守恒定律。当结果表明想法与能量守恒定律确实吻合时，海森堡非常兴奋。因为条件的限制，加上身体的疲乏，他犯了一些数学上的错误。可他锲而不舍，夜里照样工作，最后终于在凌晨 3 点把错误都找了出来，解决了所有问题。

> 一开始我很害怕。我有种感觉，就好像是透过原子现象的表面，看到了异常美丽的原子内部。一想到现在必须找出大自然如此慷慨地呈现在我面前的如此之多的数学结构，我就感到几乎眼花缭乱了。激动的心情使我难以入睡。所以在新的一天开始的时候，我来到了岛的南角，渴望登上伸出海面之上的岩石。现在，没费什么力气我就登上来了，等着太阳升起。[10]

六月底，海森堡回到了哥廷根，而且很快就计划去剑桥讲学。他用几天时间匆忙写出了一篇论文，题目是“运动学和力学关系的量子力学重新解释”（On the Quantum-Mechanical Reinterpretation of Kinematic and Mechanical Relations）。[11]“重新解释”（reinterpretation，umdeutung）一词表明了海森堡的大胆——这是一种研究原子物理学的新方法。在摘要中，海森堡大胆地称论文的目的是“为纯粹建立在‘原则上可观测量’间的关系之上的理论量子力学打下基础”。他继续写道，我们还不能从实验信息出发，“把电子与空间的点联系起来”，“在这种情况下，很显然对目前不可观测量不能抱有任何希望，如电子的位置和周期”。海森堡当时是在黑暗中摸索，实际上，位置和动量可以分别测量到任意准确的程度，只是二者不能同时被准确测定。论文指出了如何编辑与状态间的过渡相关联的振幅和频率表（海森堡把这些表称作“量子理论量”），以及这些表如何能与一种新型的微积分联系起来。海森堡将这种新型的微积分称为“量子力学关系”。

这篇论文把物理量与一种表联系了起来。其中，表的行和列分别标着“允许的”量子态。玻尔在他的那篇具有开创性意义的论文中假设了氢光谱的量子态。这种做法以前就有（如爱因斯坦的 A 和 B 系数是标着两种状态的“表”）。不过，海森堡却把这一思想应用到了一组更加基本的量上，并由此发现了两个表的“乘法”规则，得到了与经典力学中的公式相类似的公式。这一点是全新的，它开创了一条研究量子力学的新

途径，跳出了之前玻尔和其他人研究“量子力学”时的局限。

之后，海森堡意外地遇到了困难。他发明的表和表的乘法法则遵守一种新的代数。这种代数很早之前就被数学家发现了，但大多数物理学家并不熟悉，连海森堡本人都包括在内。最出人意料的是，这一法则并不遵守“乘法交换律”。根据乘法交换律，两数相乘的顺序并不影响乘积，即 $ab=ba$。海森堡用新微积分将某个量子理论表（不妨称作 A）与另一个表（B）相乘时，所得乘积与相乘顺序有关，即 $AB \neq BA$。海森堡后来说：“这点令我非常不快。”为了把它从理论中除去，海森堡使出了浑身解数，但没有成功。[12]“我感到这是整个方案中唯一的困难。这一困难如能得以解决，我会高兴得睡不着觉。”面对这一可能会威胁到新发明的困难，海森堡采取了与很多人一样的做法：掩盖真相。对此，他只在一个句子中稍提了一下——“在经典理论中 AB 总是等于 BA，但在量力理论中却并不一定是这样”，在介绍了不会出现这种困难的情形之后，他就转移话题。海森堡说自己的论文带有一种该领域早期论文中常有的“免责声明”。就量子力学来说，不知论文是“令人满意的”还是“过于粗糙的”。他称答案尚需“更深入的数学研究”。[13]

7 月 9 日，海森堡在完成论文后，将论文的复本给了波恩，请导师帮忙看看是否值得发表，同时问问自己能否研究那个在他看来似乎有些棘手，甚至有些怪异的基本思想。波恩答应了，不过因为讲了一学期的课，加上和另一个助手约尔当一直在做研究，他深感疲惫，好几天没去理会海森堡的论文。

波恩读海森堡的论文是在他走之后。这篇论文给波恩留下了深刻印象。他将论文交给了《物理期刊》，并在 7 月 15 日写了一封信给爱因斯坦，说海森堡的工作看起来“非常不可思议，但确实是正确的，而且意义深远”。[14]不过波恩对海森堡的表和表的数学相乘法则有一丝不安。这法则看上去是那么地熟悉。一周的寝食难安过后，波恩突然想到上学时在数学课上曾见过这种特殊的结构。他的这位无畏的年轻助手在不知情

的情况下居然独立地提出了这一结构。这些表被数学家们称作矩阵，表的行和列中排列着数字(或变量),而海森堡的表中元素数字有无限多个。实际上，海森堡提出的有趣的量子力学关系是数学家发现矩阵“乘法”的最自然的方式。

波恩高兴坏了。矩阵数学为他提供了一种研究海森堡的工作并使之系统化的框架。他知道矩阵是不能对易的——矩阵的乘积与矩阵相乘的顺序有关。这样一来，海森堡的难题也就得到了解释。例如，为什么动量矩阵 $\boldsymbol{p}$ 和位置矩阵 $\boldsymbol{q}$ 不能对易，矩阵 $\boldsymbol{pq}$ 与 $\boldsymbol{qp}$ 是不同的（习惯上，物理学家常常用黑斜体符号表示矩阵）。不过这还不是全部。这对变量（称为正则共轭变量）虽然不能对易，但可用一种特殊的方法使其交换位置。$\boldsymbol{pq}$ 和 $\boldsymbol{qp}$ 之差似乎是与普朗克常数成比例的特定矩阵：$\boldsymbol{pq}-\boldsymbol{qp}=\boldsymbol{I}h/2\pi i$，其中 $\boldsymbol{I}$ 是单位矩阵，即对角线上的元素都是“1”，对角线以外的其他元素都是“0”。波恩后来写道：“我看到这个结果时非常激动，就像一个水手经过长久的航行，远远地看到了期待已久的陆地时的心情。可惜海森堡当时没在，真是遗憾。”[15]

几天后的 7 月 19 日，波恩在火车上遇见了泡利。波恩激动地向泡利解释如何能将海森堡的论文翻译成矩阵语言，并问这位自己以前的助手是否想合作研究这一课题。可泡利却没放在心上，反而讽刺波恩要用“无用的数学”和“繁琐复杂的形式”“糟蹋海森堡的物理思想”。(历史学家认为这个评论很幽默，因为在这里海森堡的思想是形式化的，甚至比传统的矩阵分析还繁琐。）第二天，也就是 7 月 20 日，波恩拜访了约尔当。约尔当具有深厚的矩阵数学功底。不出几天时间，两人就弄清楚了如何从海森堡的工作推导出 $\boldsymbol{pq}-\boldsymbol{qp}=\boldsymbol{I}h/2\pi i$ 这一关系。这一次，波恩又是惊奇不已：“我永远忘不了海森堡的量子状态思想被成功地浓缩到神秘的 $\boldsymbol{pq}-\boldsymbol{qp}=\boldsymbol{I}h/2\pi i$ 方程中时，自己的紧张心情。这个方程是新力学的核心，之后我们发现它隐含着不确定关系。”[16]

9 月底，他们寄出去一篇论文，题目是“论量子力学”（On Quantum

Mechanics）。论文中包含了海森堡所期望的“深入的数学研究”，这也是首次用公式表示矩阵力学。论文所采用的数学是人们所不熟悉的，许多物理学家必须在研读过矩阵之后才能理解论文的内容。数学方法也比较笨拙，不过应付有限的问题，从头一直算到尾还是没有问题的。两位作者把论文的复本寄给了海森堡，此时他已经离开剑桥，正在哥本哈根。他把论文拿给玻尔看，说：“瞧，波恩给我寄来一篇论文，可我却一点也看不懂。全是矩阵，都不知道是干什么用的。”[17]不过，海森堡在复习了矩阵知识后，他也分享了波恩和约尔当的激动心情。7 月 18 日，海森堡写了一封信给泡利，信中说波恩的智慧思想 $\boldsymbol{pq}-\boldsymbol{qp}=\boldsymbol{I}h/2\pi i$ 是新力学的基础。海森堡、波恩和约尔当三人于是开始了狂热的通信。海森堡也终止了在哥本哈根的逗留，返回哥廷根，以便三人能够完成另一篇论文的工作——在波恩早先计划好的 10 月美国之行前，对波恩–约尔当论文的结果加以推广。[18]

三人合作的结果就是由波恩、海森堡和约尔当合写的题为“论量子力学 II”（On Quantum Mechanics II）的论文。物理史学家把这篇论文称作“三人论文”（the three-man paper）。论文的核心是所谓的“量子力学基本关系”，也就是那个奇怪的方程 $\boldsymbol{pq}-\boldsymbol{qp}=\boldsymbol{I}h/2\pi i$。这篇论文是物理学历史上的一个里程碑，它第一次描述了量子领域的图景。几乎同时，泡利也发表了一篇论文。在论文中，他成功地将矩阵力学应用到了氢原子的例子上（克服了相当大的困难）。

然而，除矩阵力学的发明人外，几乎没什么人能够认识到矩阵力学的重要意义。这其中的原因在于，人们对矩阵力学价值的认同遭遇了几个障碍。障碍之一就是矩阵力学本身的复杂性：矩阵力学本身并没有那么难，可是海森堡在应用它的时候却看似异常复杂。大多数物理学家在掌握的过程中只能不加怀疑地接受。乔治·乌伦贝克（George Uhlenbeck）的反应就很典型。他当时是莱顿大学的一名学生。乌伦贝克后来说：“所有一切都变成了这些必须要解出来的难计其数的方程，所以没人知道到

底如何去做。”[19]其他一些人则因下述事实望而却步，即矩阵力学故意不给出原子力学的图景，它的基本项（矩阵）严格来说是没有意义的，只是形式上的象征物。[20]还有一些人对矩阵力学不能解释微观世界到宏观世界的过渡——没有空间和时间且无法看到的世界与人类生活在的、想象力所熟悉的时空背景之间的过渡，深感苦恼。历史学家马拉·贝勒（Mara Beller）曾写道：许多科学家于是采取了“观望”态度，甚至连波恩、海森堡和约尔当等人都把它当成是朝着正确理论迈出的并不完美的第一步。[21]

然而，三人的论文在1926年2月问世后不久，作者们就邂逅了一位讨厌的同伴。

矩阵力学与波动力学

薛定谔有关波动力学的第一篇和第二篇论文分别于1926年3月和4月发表于《物理学年鉴》。波动力学与矩阵力学的研究对象相同，不过物理学家发现前者更易懂。波动力学中不存在矩阵力学中的那些障碍。第一，波动力学所涉及的数学是经典物理学家基本训练内容的一部分。从高中时代起，他们就一直在使用、求解波动方程。第二，波动方程是直观的。物理学家能看到水、声音、光波。这些波的属性（频率、振幅和波长）沿着波光滑、连续地传播。物理学家自己通过训练，可以看到波的其他性质，如波节和干涉等。不过这些都是 Ψ 函数中的小问题。波函数存在于多维“构形空间”之中——对系统中的所有粒子而言，波函数是三维的。即便如此，这些粒子看上去似乎也是直观的，就像某个物体在空间中运动或者被局限在原子内部像驻波一样保持“静止”。第三，波动力学提供了一种从微观世界到宏观世界的过渡的自然描述方法。就像薛定谔在当年发表的第三篇论文指出的：好比粒子一样，并且沿着经典路径运动的波组或者波包。这些经典路径是与 Ψ 函数的相前垂直的光线。[22]

大多数物理学家并未对波动力学感到吃惊。然而，普朗克对此确实惊奇不已，而爱因斯坦更是狂喜。美国物理学家卡尔·达柔（Karl Darrow）曾报道说，波动力学"捕捉到了物理学的世界"，因为它预示着"长久以来令人们备感困惑、无法抑制的那种重返经典物理学的期望"（人们更加习惯的连续传播函数）将得以实现。[23]大量论文开始采用薛定谔的方法来解决原子领域的问题。不过，哥廷根物理学家的拥护者却有不同的反应：海森堡说波动力学"看起来太好了，都不可能是真的"；狄拉克的反应更是充满"敌意"；泡利则把波动力学说成是"疯狂的"。[24]然而，这些人很快就拜倒在波动力学的石榴裙之下。泡利用矩阵力学，辛辛苦苦地算出氢光谱理论后不久就发现，用波动力学去算氢光谱理论更加容易。波恩在给薛定谔的信中写道，他在读了矩阵力学的第一篇论文之后，非常激动，甚至都想"背弃连续物理学，背弃经典物理学脆弱、清晰的概念基础"。[25]不过他的热情很快就降温了。

起初，冲突集中在两种方法的科学价值的争论上：哪一种方法的效果更好？第一个检测对象就是氢原子。这个问题泡利已经分别用两种方法解决了。氢原子就好比是原子物理学家的果蝇或者实验鼠。所有模型都需要解决的第一个问题就是它：因为氢原子已经用旧量子论成功分析，并且结论与实验结果非常吻合，因此这也意味着可对公式加以比较。另一个检测是要解释量子世界与经典世界之间的过渡，也就是如何从原子世界过渡到人们所熟悉的宏观世界。薛定谔已经表明波动力学可以给出答案，但矩阵力学目前对此尚无能为力。不过，另一个重要问题是如何处理原子世界中的碰撞问题，这需要说明系统随时间的变化情况。

"究竟哪种方法具有更重要的科学价值"的问题很快就有了答案。1926 年 5 月，在他的第四篇论文中，薛定谔证明了两种方法在数学上是相同的。[26]泡利也得到了相同的结论。虽然尚不清楚如何处理所有的检测，但是两种方式在数学上是等价的这一事实已经表明，两者在数学上的重要性是旗鼓相当的。不过，从科学上说，波动力学对波谱连续部分

的分析是必需的，因此它的用途要比矩阵力学广。

海森堡的自传作家戴维·卡西迪（David Cassidy）指出，尽管如此，这一结论却使冲突再起，令各方也开始严肃对待。随着数学上的等价性这一问题的解决，支持者们现在可以针对理论的物理解释进行自由辩论。两方针锋相对，不过薛定谔对波动力学的解释至少还是给人以希望的。它是这样描述的：随着时间和空间的展开，原子世界乃是自连续过程编织交错而来；这些过程又恰是事件看似不连续的原因。相反，矩阵力学所描述的原子世界中没有连续过程和偶然联系，与时间和空间也没有关联，并认为原子世界是一个人类无论如何也想象不出的世界。与科学价值上的冲突相比，该冲突容易确定，不过也更倚赖于人的情绪。它反映出的是对手对物理学的认识，对世界的认识，以及对人类和世界之间的最根本关系的认识。

不过，正像贝勒所指出的："在有关解释的争论的起始阶段，没人有一个能够明确说清楚的位置，更不用说了解'真相'了。"[27]现在，这种再起的冲突迫使支持者们就各自理论的物理解释进行辩论。薛定谔也只能辩论道：在最基本的层面上，世界是由连续性组成的；要描述它，并不需要海森堡那套蹩脚的正规方法。薛定谔还必须要解释波包如何能结合在一起，详细阐述 Ψ 函数的意义，说明量子现象的非连续性为何是源于连续的波动过程的。最终，薛定谔只能承认波存在于多维构形空间中。海森堡和他的支持者只能辩说整个世界全是非连续的，用除此之外的其他方式来呈现都会误导他人。他们还必须给出一种把矩阵力学的正规符号项与人们所熟悉的性质联系起来的方法，说明为何在到了波动力学可直观理解的程度后，该方法就是错误的。对双方间的这种不一致，各自的支持者们已经习惯了。贝勒表明了各方是如何"做手脚"，把对方理论的内容加到自己的当中，并使其发挥作用的。但是，这种新的、强烈的冲突却为不确定性原理和哥本哈根诠释的提出铺平了道路。[28]

薛定谔已经开始在论文中向对手发动攻击，证明两种方法的同一性。

两种方法的确是相同的，不过他说他却因矩阵力学“极其困难”的数学方法和无法直观理解的特性而“失去信心，或者说因此就被拒之门外了”。[29]他后来还说，如果人们“压抑直觉”，“只用跃迁概率和能级等抽象观点的话”，那么要解决跃迁之类的原子领域的问题就会变得“极其困难”。[30]薛定谔在给维恩的信中写道：公开表示把物理学局限到客观测量之必要性上，“无非是为了掩盖人们无法猜想出真实图像这一事实”。[31]

海森堡说的话至少算得上是尖刻的。他把波动力学描述成“恶心的垃圾”。他声称，波动力学要是到了能够直观理解的程度，它就是错误的，使用矩阵力学的物理学家较不易被迷惑，从而能够更深刻地洞察自然。[32]波恩曾说：“数学比直觉更可靠。”[33]

双方之间的冲突很快就演变为面对面的对抗。1926 年 7 月，薛定谔和海森堡在慕尼黑的一个会议上首次见面。薛定谔此时拥有更多的支持者。他做了两个关于波动力学的报告。海森堡站在最后，反驳说：没有一个依赖于连续过程的理论能够解释量子现象的非连续性，如普朗克辐射和康普顿效应。不过听众似乎还是站在薛定谔一边，似乎没什么人受了海森堡的影响。海森堡也感觉到一丝挫败感。之后，他去了哥本哈根，呆了几个月的时间，与玻尔一同工作。可两人的意见也还是不一致。玻尔认为必须采用经典概念对实验现象进行描述，海森堡却不这样认为。不过他们却就如下问题达成了一致：为什么量子不连续性意味着无法定义空间和时间，也就是说量子世界既不能用包含连续函数的理论来表示，也无法通过人脑（想象力依赖于时空背景）想象出来。

薛定谔和矩阵力学派之间的下一次较量发生在三个月之后。1926 年 10 月，玻尔邀请薛定谔到哥本哈根参观。哥本哈根是矩阵力学的阵地（虽然哥本哈根学派也已经开始将某些波动力学理论作为工具使用）。薛定谔学术态度理智诚实，似乎又赢得了广泛的认可，自然很高兴能去参观对手的总部。不过，他对后来发生的事情却完全没有准备。玻尔在火车站接到了薛定谔，他立即开始向薛定谔介绍工作的进展，与薛定谔没日没

夜地辩论了好几天。玻尔安排薛定谔住在自己家里，哪怕一分钟都不想放过。海森堡后来回忆道：

> 玻尔是一位非常体贴和热情的人。但是，在这类他认为具有重要意义的认识论问题上，玻尔却坚持所有论断都必须是完全明确的。这种坚持甚至演变为狂热的、恐怖的坚韧。辩论进行了几个小时，玻尔还是不肯放弃，直到薛定谔承认自己的解释还是不够，甚至不能解释普朗克定律。也许是因为辩论过度劳累的关系，几天后薛定谔病倒了，躺在玻尔家的床上。即便如此，玻尔也还是不肯从薛定谔的床边离开。他会一遍又一遍地说："可是薛定谔，你至少必须承认……"有一次，薛定谔终于爆发了："如果你还要坚持你的狗屁量子跃迁的话，那我可真该后悔当初踏入了原子论领域。"[34]

有海森堡在身边，玻尔说服了薛定谔作出（临时的）让步。不过这一让步并没能持续多长时间，薛定谔很快就开始撰写波动力学的论文。1926 年 11 月，薛定谔把他的六篇波动力学的重要论文（发表在《物理学年鉴》上的四篇系列文章"特征值问题的量子化"以及分别关于边界问题以及波动力学和矩阵力学的同一性的两篇论文）整理后出版成书。

此时，玻尔对波动力学作出了新的解释。他想要理解电子和原子之间的碰撞，于是仔细检查薛定谔的结论，即 Ψ 函数指的是电子的电荷密度。他发现这一结论讲不通，并做出结论：Ψ 函数说明的并不是事件的状态，而是事件的概率。之后，泡利写了一封信给海森堡，提出 Ψ^2 表示粒子在特定位置上出现的概率，而不是状态的概率。这样就能在一定程度上恢复时空背景和便于直观理解。电子从某个位置移动到另一个位置的轨道或者路径能否直观想象并不是必须的，但是不管电子是通过什么路径到达的，它总会具有一定的位置。[35]所以，经典性质的确存在，而且可以精确测定。不过，那个奇怪的想法仍然是摆脱不了的，即薛定谔

说的在空间流动的奇怪函数并不是实际存在的物体，而只是实际物体在那一点能被发现的概率。波恩后来评论说："我们已经习惯了作统计上的考虑，所以更深一层于我们而言并无太重要的意义。"[36]

在 10 月 19 日提出波函数解释的那封信中，泡利还注意到了恼人的 ***pq*−*qp*** 问题的深层含义。海森堡曾辩称：这两个共轭变量（不可对易项）指的都不是位置或动量等可以共同准确测量的经典变量。泡利说两者之一可能是经典变量——但如果是这样，另一个就只能是概率了。这使不可对易性变得更加奇怪。泡利告诉海森堡："我对这方面的物理学从头到尾一点都不清楚。"他曾困惑地说："我的第一个问题是，为什么只能用任意精度描述 ***p***，而不能以任意精度同时准确描述 ***p*** 和 ***q***？""你可以用 ***p*** 眼或者 ***q*** 眼看世界，可同时睁开两只眼看，却会出错。"[37]这意味着什么？

海森堡的回复有点迟，主要是因为泡利的信在同事间传阅，很难拿到。10 月 28 日，海森堡终于寄出了回复。他仍旧没有接受对直观化和经典变量的隐含的恢复，并认为波恩"极其教条"的观点只是"诸多可能的解释中的一个"，而不予理睬。他还继续坚持认为 $\boldsymbol{pq}-\boldsymbol{qp}=\boldsymbol{I}h/2\pi i$ 这一关系表明了单独的 ***p*** 或者 ***q*** 都是没有意义的。"总的来说，我希望答案是以下类型的（不过他并没有到处宣扬）：时间和空间只具有统计学上的概念，就如同气体的温度、压力等。我的观点是，在谈及单个的粒子时，空间和时间概念都是没有意义的。粒子数目越大，时间和空间的概念就越有意义。我一直想把它继续向前推进，可到目前还没有成功。"

几个星期后的 11 月 15 号，海森堡向泡利展示了一个看似结论性的论证：为什么量子世界使单独的 ***p*** 或者 ***q*** 失去了意义。[38]不妨假定在某个特定的点上有一个物体，比如电子。电子的速度是由它在该点与离该点无穷远的点之间连续运动时的速度所决定的。但是，如果时空是非连续的，电子就会从一个状态飞到另一个状态。这样，从定义上来说电子就没有速度了！一个星期后，海森堡回来了，依旧对这一问题深深着迷。[39]

世界是不连续的，所以“c 数”（经典数）表明人们对发生的事情了解得太多了。“对于‘波’和‘粒子’两词的意思，反而弄不清楚了。”

约尔当现在也开始对海森堡提出挑战。实际上，约尔当一直扮演着海森堡的无力假设（postulate-of-impotence）的反对者的角色。海森堡的假设是：单个的电子不能具有位置和动量。那么，阻碍实验者对位置和动量进行测量的因素是什么呢？观测仪器是由原子构成的。由于热运动，原子在室温下会发生振动，从而使准确度具有一个实际限度。如果能在绝对零度下搭建起一台仪器用于测量，那么情况会如何呢？或者为了达到同样的效果，采用 α 粒子（可以忽略振动，并可以记录路径）之类的高能探头效果又会如何？

波恩和泡利考虑固定一个变量在理论上的可能性，并注意到另一变量只能说具有一定的概率。现在，约尔当指出了物理学家测量原来被认为无法测定的量的实验条件。这个量就是“在特定位置发现电子的概率”。从理论上说，不是不能观测，只是在实验上比较困难。

约尔当的论文困扰着海森堡。[40]论文发表后的那天，也就是 1927 年 2 月 5 日，海森堡在给泡利的信中写道他发现约尔当的论文“虽然很好，但有很多处是不准确的”。这是因为约尔当仍然认为“在特定位置上发现电子的概率”之类的叙述是没有意义的。但是，如果单个原子的时间和位置等在实验上有意义，那么它们在理论上也一定是有意义的。而如果理论上有意义，约尔当的方法就是错误的。

在所有的讨论中，人们从未怀疑过数学的正确性。存在争议的是解释，甚至是解释的本质。玻尔想从解释中获取的较海森堡为多，而这两人又较薛定谔为多。

海森堡仍在哥本哈根玻尔的研究所工作，住在玻尔的哥哥赫拉德的顶层公寓中。晚饭后，玻尔就会拿着烟袋来到海森堡的住处，对量子力学的现状进行辩论，两人一直能谈到凌晨。这种费力的谈话开始慢慢消磨两人的关系，双方都变得容易动怒。有感于此，玻尔外出去滑雪了。

玻尔走后，一天晚上海森堡在玻尔研究所后面的费莱德公园（Faelled Park）散步。他从理论和实验上都仔细考虑了 $\boldsymbol{p}$ 和 $\boldsymbol{q}$，也考虑了约尔当的显微镜。像往常一样，他坚信约尔当的例子中某个地方一定是错的。约尔当使海森堡变得理智，令他不再一心去思考理论意义。他必须不再从哲学的观点去思考概念，而是从操作的层面思考实验者要做些什么。假定在绝对零度下对电子进行观察，这也等于是说从粒子表面弹射出一个光子，并在仪器的透镜系统中捕捉到它。但是光子的弹射将会对电子的位置产生扰动。为了避免扰动，就要采用能量较低的光子。而光子的波长越长，对位置的测量就越不准确！海森堡激动地发现，这个问题之所以会出现，是因为仪器和待测物之间的相互作用，也就是采用的观测工具与被观测系统之间的相互作用。

开始新纪元

接下来，海森堡所做的与他一激动时的惯常做法就没有二异了：他写了一封信给泡利。这封于2月23日写就的信不是一般地长，有40页之多。海森堡的这种思想上的转变其实是由约尔当的论文所开创的。他描述了几个测量 $\boldsymbol{p}$ 和 $\boldsymbol{q}$ 的理想实验，所以从一开始这种转变就很明显。后来他写道："人们总能发现，所有理想实验都具有这一特点：如果量 p 的准确度被限定在 p 的平均误差以内的话，那么 q 的准确度同时就只能限定在 $q_1 \approx h/p_1$ 的范围之内。"

这就是不确定性原理。像其他许多方程一样，不确定性原理最初出现时的形式也并不是现在人们熟知的形式。如今，这一原理常被写成不等式的形式：$\Delta x \Delta p \geqslant \dfrac{h}{2}$。

不确定性原理是一个概念上的突破。波恩、泡利和约尔当都考虑过共轭变量之一能够准确测定，而另一个变量却变成了概率的情形。海森堡现在说明：上述情形都是极端情况，在这些极端情况之间还有其他一

些两个值都不确定的情况。一定程度的不确定度是不可避免要存在的。比如，如果说电子位置的不确定度（Δx）很小，那么动量的不确定度（Δp）一定要很大，才能保证二者的乘积 $\Delta x \times \Delta p$ 与 h 在一个数量级上。如果电子位置的测量精度的不确定度很小，那么响应的电子动量的不确定度就会很大。海森堡告诉泡利这是 $\boldsymbol{pq}-\boldsymbol{qp}=\boldsymbol{I}h/2\pi i$ 的直接结果。由此，$\boldsymbol{pq}-\boldsymbol{qp}=\boldsymbol{I}h/2\pi i$ 的解释最后似乎终于明朗起来。海森堡将粒子重新放回到时空舞台，至少是暂时放上去了，但却赋予它们明显的非经典性质。

海森堡很快就写了一篇论文，表述自己的思想——“量子动力学与量子力学的直观理解”（The Visualizable anschaulich Content of Quantum Kinematics and Mechanics）。这篇论文以向受过经典物理学训练的物理学家解释量子力学如何能通过经典项得以直观理解为出发点。为了达成这一目标，论文的第一句就重新定义了“可直观理解性”：“我们相信，如果在所有简单情况下都能看到定量的实验结果，以及检查过理论的应用中不含有内在矛盾后，我们就理解了理论的可直观理解性。”这些定义既抓住了核心，又便于使用，于是海森堡最终就可以使自己的理论与这些定义相适应。不过海森堡又说量子力学看上去很难与该定义相适应，原因在于：只要 $\boldsymbol{pq}-\boldsymbol{qp}=\boldsymbol{I}h/2\pi i$ 成立，所谓的位置和速度等概念就是不清楚的，需要通过指定实验条件来澄清事实。因此假设我们观察一台显微镜下用光照明的电子。因为电子很小，必须要用高能光：γ 射线。但如果用高能光去照射微小物体，又会产生康普顿效应，光子与微小的电子发生碰撞，会以非连续的方式突然将电子撞到一边。海森堡这样写道：

> 采用的波长越短（即位置测得越准确），（非连续性）变化量就越大。电子的位置在一瞬间被确定后，它的动量就只能准确到与该非连续性变化对应的几个量级了。从而，位置测得越准确，动量就越难以被准确测定，反之亦然。在此种情形下，我们看到了方程 $\boldsymbol{pq}-\boldsymbol{qp}=-ih$ 的直接物理解释。

海森堡对在方程中使用单位矩阵 $\boldsymbol{I}$ 向来是漫不经心，甚至常常在论文中将其省略。为了对上述解释进行定量，海森堡继续说道：

> 假定 q_1 是测定 q 的准确度（例如 q_1 是 q 的平均误差），q 在这里是光的波长。假定 p_1 是测定 p 的准确度，也就是在康普顿效应下 p 的非连续变化。于是，根据康普顿效应的基本定律，p_1 和 q_1 有如下关系：
>
> $$p_1q_1 \sim h$$

这时出现了一个奇怪的问题，这个问题的重要性直到最近才被约翰·马伯格（John H. Marburger. III）发现。海森堡接着说，该方程是“方程 $\boldsymbol{pq}-\boldsymbol{qp}=-ih$ 在数学上的直接结果”。不过海森堡并没能说明这一点。在他的论文中，也找不出不确定性原理的推导。虽然这一原理为海森堡和玻尔两人所接受，不过很明显它只是一个猜想，没人自讨无趣地去证明它。玻尔所引用的该原理的首个证明是有问题的。[41]

这篇“直观化”的论文不再像两年前的那篇“重新解释”的论文那么激进。它并不说电子没有位置或速度，而只是说它不能同时具有确定的位置和速度，使位置或速度具有确定值。海森堡还原了足够的直观化能力，称“量子力学不应再被认为是抽象的、不能直观想象的。”他引用了薛定谔认为矩阵力学“令人恶心和战栗”的说法，转而提出真正的敌人是薛定谔对直观理解的误解。这是对薛定谔的致命打击（coup de grâce）。原子世界是可直观想象的，但是能够直观想象的显然并不是经典的。不过，在仔细读过后，反而搞不清楚海森堡是否对可视化还有一点坚定的立场。贝勒写道：“海森堡假定了世界的经典图像，为的却是反驳这一经典图像。”[42]

论文完成后，海森堡写了一封信给约尔当。信中说：他感到“非常非常高兴”，经历了整整一年时间的悬而未决，自己现在觉得“脚下的土地都是非连续性的”[43]。泡利听后非常激动。“海森堡所说的像是具有开

启新纪元（Morgenröte einer Neuzeit）般的重要意义。”[44]

然而，新纪元的开启也面临着重重困难。玻尔回来后，海森堡给他看了论文。玻尔指出了论文中的几处明显错误。玻尔指出，即便是在原子世界中，能量和动量也是守恒的。如果用光子撞击电子，那么只要能捕捉到光子，就能得出电子的动量，消除不确定性。不过玻尔接着说，海森堡的思想仍然是正确的，只是正确的原因在于粒子的波动本质。不过这些反冲粒子的动量是无法准确测定的——即便是用电子替代光子也无法做到。这其中的原因在于粒子是按照薛定谔方程所描述的像波一样散开的。这也是为什么要采用显微镜将其聚焦。但这也意味着承认薛定谔的波在理论中发挥了重要的作用。不过玻尔和海森堡之间的对话很快就恶化了，两人各自坚持自己根深蒂固的立场，谁也不肯让步：玻尔认为需要波，海森堡认为可以没有波。玻尔告诉海森堡不要把论文发表，令海森堡很沮丧，大哭一场。[45]不过正像贝勒指出的，海森堡大哭一方面是因为玻尔的无情，另一方面也是因为他本人的固执。

海森堡无视玻尔的建议，拒绝撤稿，甚至拒绝对论文加以修改。他只是附加了一个短评，题目是“证明之补充”（Addition in Proof）。短评中说：“玻尔已经提请我注意本论文有几处讨论忽略了重要的几点。”不过他并没有对这些被忽略的点进行修改。

一连几个月，玻尔和海森堡仍旧未能就量子力学的解释达成一致。根据爱因斯坦的记录，两人对数学的正确性都没有异议，也都认为一定要用数学来指导解释。不过关于如何用数学来指导，玻尔的想法更好。海森堡认为可以采用矩阵语言或者波动语言，而玻尔认为两者都是需要的。海森堡的立场其实是柏拉图式的：他想说的是仅用数学就足以描述原子世界。玻尔的立场则是康德式的：他认为自然界迫使人们按照由时空舞台衍生出的某些特定（经典）范畴和图式来感受和想象。就像马伯格所说的：现实是宏观现象。这些范畴和图式适用于宏观事件以及为事件提供理论基础的经典物理学。它们并不适用于微观事件——如果硬要

假设它们有效，并应用于微观事件，就会产生所谓的宏观谬误（macroscopic fallacy）。我们也无法在思考和想象时避免这些经典图式。由此玻尔推论说：人们在思考微观世界时，被迫要依赖于传统的范畴和图式（如位置和动量）；然而人们却在以重叠的、非经典的方式使用这些范畴和图式，就像在“互补”对中使用时那样。我们必须要抛弃这种观念：适用于宏观世界中可感觉到的现象的概念和图式，是与微观世界中的事物相对应的。这样一来，玻尔的康德式的方式就充当了在量子理论和“真实”现象的世界之间建立起本体论联系的角色。至此，上述观念已经是奇怪得不能再奇怪了。“一般物理意义下的独立实在既不能归于现象，也不能归于观测手段。”[46]

不过，1927 年末，玻尔计划要去一趟美国。他和海森堡都很希望能在走之前完成对量子理论的解释。两人主要基于玻尔的主张，达成了一致。当年 9 月，在科摩湖召开的亚历山德罗・伏打（Alessandro Volta）逝世一百周年纪念会上，两人的“休战”被公开。玻尔在会上作了报告，指出了波动力学与矩阵力学之间尴尬的妥协。海森堡最后也站起来表示赞同。玻尔说：实际上波和粒子是在谈及原子世界时常用的方式。两种方式都不完全准确，二者的应用范围有限，且相互交叠。他宣称：波和粒子是在谈及人们并不直接了解的事物时所采用的两种“互补”的方式。正像玻尔曾说过的那样：“其实没有量子世界，只有抽象的物理描述。那种认为物理学的任务就是去发现自然是什么的想法是错误的。物理学关心的是如何描述自然。”[47]

这就是后来成为所谓的量子力学的哥本哈根诠释的起源。这一诠释并未能赢得广泛的认同。爱因斯坦说它“不可靠”，还说“海森堡–玻尔的安逸哲学（或说是宗教？）在设计上真是精妙无比。眼下看来，每个真正相信此哲学的人，它都送一个柔软的枕头，一旦躺在上面就很难叫醒了”[48]。几年后，爱因斯坦在同波里斯・波多尔斯基（Boris Podolsky）和纳森・罗森（Nathan Rosen）一起完成的完整理论中写道：“物理实在

的所有元素在物理理论中都要有一个对应物。”爱因斯坦还尝试证明：量子力学的不完整性其实是一个瑕疵，表明还有更多所谓的隐变量有待于去发现，这些隐变量的发现将会使量子力学公式直接指向现实世界。不过爱因斯坦没能成功证明这一点。这一论点爱因斯坦坚持了多年。不过玻尔表示了反对，认为位置和动量本质上是经典概念，不适用于微观世界的事件。如果硬要应用到微观世界，也只能通过不精确的，严格地说是不准确的方式。

哥本哈根诠释是一个清晰的、富有逻辑性的解释，而且在所有符合理论和实验约束的解释当中似乎是最简单的一个。它指的是：在超越或者靠近微观世界的某处潜藏着人们无法看到的规律；通过具有宏观行为的物体的组合或者排列，就能看到规律。这个解释让所有人都感到不舒服。不过这是正常的心理现象，不能因此就赞同或反对它。



茶　歇

神秘莫测的量子论

“中间类型的实在”这一思想是必须要付出的代价。

——海森堡

1929 年，也就是不确定性原理出现后的 2 年，哈佛大学一位叫做珀西·布里奇曼（Percy Bridgman）的物理学家在《哈珀杂志》（*Harper's Magazine*）上发表了一篇关于不确定性原理的意义的论文。他说不确定性原理哪怕对公众也有着深远的意义。“该原理立竿见影的效果将会导致

真正的天马行空式思维方式的出现。”布里奇曼继续说，科学家以外的人倾向于从不确定性原理得出结论：不确定性原理叙述的并不是“意义的结束”，而是“还存在着超越科学家知识范围的其他东西”。在一篇颇具预言性质的文章中，布里奇曼写道：

> 这种想象（科学家已证明他们无法洞察）成为了所有神秘主义者和梦想家的游乐场。这一领域的存在导致了漫无边际的各种解释的出现。它将成为灵魂的本质，死者的灵魂填在其中，上帝潜伏在它的阴影中，重要过程的原理在这里可以找到，它将成为心灵感应的媒介。一个小组将通过因果物理定律的失效，找到一直以来悬而未决的意志自由问题的答案。而另一方面，无神论者则找到了“机会掌控着宇宙”这一论点成立的理由。[49]

80 年后的今天，我们发现布里奇曼是对的：他的所有观点的确都是很超前的。布里奇曼继续站在积极的一面，认为最终可以提出“新的教育方法”，教给人们如何去重塑在“日常生活的极限情况”下进行思维所需的“思维习惯”。布里奇曼认为最终的结果将是有益的。

由于思维要符合现实，因此人类理解和征服世界的步伐将会加快。我大胆地猜想，这对人的个性最终会产生良好的影响。虽然一般人会消极应对，可是需要有一些具有勇敢的高贵品质的人来面对类似的情形。最终，当人类充分分享了知识树的果实后，第一个伊甸园和最后一个伊甸园就会产生这种不同：人不会成为上帝，他永远都是谦卑的。

80 年后，我们还在对勇敢的高贵品质孜孜以求，同时也依旧是谦卑的。不过我们也在研究如何从物理上解释量子力学，以及量子力学是如何与其他人们更加熟悉和可直观认识的世界特征联系在一起的。

在所有的量子力学创立者中，玻尔是最强硬地坚持应该在经典概念

的通俗语言框架下，来对量子世界进行充分的表达的。在弗雷恩所写的剧本《哥本哈根》中，“玻尔”曾说过：“最终，我们必须要能向玛格丽特解释它。”玛格丽特是玻尔的妻子和秘书，在舞台上代表一般人（也就是思维传统的人）。

许多物理学家在发现任务不切题或者不可能之后，对一些片面的解释也就满足了。海森堡说数学上行得通——这就够了！可玻尔却对这种托辞说不，从而使物理学家能知道自己哪里没有理解，或是掩饰了什么。玻尔自己也知道，他并没有答案。不过并不能由此就认为找不到答案。他的最接近的答案是互补学说。该学说采用一般的语言，说的是量子现象究竟是表现为波，还是表现为粒子（表面上看是矛盾的），取决于仪器的设置。要想充分抓住这一现象，这两个概念缺一不可。可正当物理学家就量子力学的“含义”问题争执不下时，学说和讨论却都消失了。

为什么会这样？很大一部分原因是到1930年时，物理学家已经在量子框架下找到了一种非常合理的描述经典概念的方式。这种方式采用一种特殊的抽象数学语言，称为希尔伯特空间（无限维空间）。在希尔伯特空间中，位置和动量的概念分别与两组不同的坐标轴（并不对齐）联系在一起，并由此得出了用普通词汇——“互补性”（complementarity）[50]来描述的情形。玻尔采用互补性的概念说明量子现象既是粒子，也是波。此时用普通词汇描述确实有点令人费解。希尔伯特空间则提供了另一种选择，和一个更加准确的框架。在该框架下，可以说量子现象既不是波，也不是粒子。但是玛格丽特并不能理解这种语言，对她来说，量子力学必须要保持它的神秘性。同时，玛格丽特也必须要尽自己所能去理解它。这就为本章开头以及布里德曼提到的巧妙解读量子力学对人类生活的意义打开了一扇门。

到底是什么使得量子力学的解读如此之困难？原因在于我们期待的完整理论没能充分地描述自然。它的描述方式具有特殊性，定义很明确。这么说是因为它提供了一个测量的理想极限模型，认为理论和实验室测

量结果之间的任何差异或偏差都来自测量设备的误差和瑕疵。如果物理学家为了获得良好的近似而略去自然的某些方面的话，那么他们教授和使用的其他理论和方程也都会有各式各样的差异或偏差。$F=ma$ 就是一个例子。为了应用方便起见，该公式把质能转换忽略掉了。我们可以把这些公式称为“无关大碍的近似或者不完整的描述性理论”。理论和现实世界中的任何差异都是认识论上的。也就是说，这些差异一定与人们对世界的认识有关，或者与人们对世界的表述与真实世界之间的差异有关。

不确定性原理的不完整性又是另一种不同意义上的。该原理是一个数学关系，是量子力学中波函数的一个统计解释特征。它没有引用任何基本的物理图像，也没有引用波、粒子以及物理实验。它的所指并不明确，只可能是探测器的计数。不过，该原理所言的差异却是世界本身之中的。这些差异不是认识论上的，而是本体论上的，它们并不是与人们的知识有关，而是与世界有关。

这很奇怪，为什么？很重要的一点是要看一下这种奇异性不是由哪些因素引起的。不确定性原理的特殊性并非是由于测定过程干扰了被测量的物体（这是涉及粒子交换的任何牛顿理论的特征），也不是由于统计的存在。相反，量子力学的这种特殊性在于量子方程的对象并非是“关于”传统意义上的真实物体或者理想物体。

在经典物理学中，量的测量值和理想值之间的偏差是用统计误差理论单独来处理的。但是，量子测量的偏差（统计分布）却是与一个单一形式系统联系在一起的。需要准确知道分布的宽度。不能对单次测量的结果进行预测。叠加后得出的结果就是世界的概率——观察的概率。因此，量子力学的波动方程的对象既不是理想物体，也不是真实物体，而是一种可以接受各种可能实验存在（experimental realizations）变成实际物体的特殊性半抽象物体。如果用针对真实世界中人们熟悉的事物的思维方式去考虑这种特殊的半抽象物体，那么它们将是不完整的。要将抽象物体带进真实世界，就要加一些东西到这些物体上。不同的测量对象

和测量手段会影响测量的结果。例如，根据所处的环境，抽象的波可能表现为波或是粒子。波和粒子在时空中是可以看到的。

这就是“中间类型的实在”（intermediate kind of reality），即海森堡所谓的要得出量子现象就必须要付出的代价。它有着剧本或者乐谱等所具有的那种不完整、半抽象式的有趣真实感——它们说到底只是为世界中的真实事物（拍摄的电影和表演的音乐）所设计的程序。真实事物需要增添背景，而背景的选定会影响到这个抽象物体。于是，牛顿没有考虑的“人类的目的和决断”（human purposes and decisions）在这里就被提出来了。

向非科学专业人士解释不确定性原理的挑战就在于得解释清楚这种新型的半抽象事物。作出这一步尝试很重要，不然公众在理解不确定性原理时，信息的丢失和曲解将依旧存在。

海森堡证明了这一点，但并非是从数学上。

注　释

[1] Anne Bogart 和 Kristin Linklater，《平衡动作》（*Balancing Acts*），《美国戏剧》（*American Theatre*），2001 年 1 月。

[2] David Cassidy，《不确定性：海森堡的一生与科学》（*Uncertainty: The Life and Science of Werner Heisenberg*，纽约：弗里曼，1992 年）。

[3] 访谈，1963 年 2 月 27 日，海森堡，《量子力学史档案》（*Achieves for the History of Quantum Mechanics*，后称 AHQP，马里兰大学帕克分校：美国物理研究所），第 22 页。这并不是当时海森堡要说的话。彼时，他正准备把所有经典概念都除掉。

[4] 海森堡，《回忆量子力学的发展》（*Erinnerungen an der Zeit die Entwicklung der Quantenmechanik*），刊于《二十世纪的理论物理学》（*Theoretical Physics in the Twentieth Century*），M. Fierz 和 V. F. Weisskopf 编，纽约：Interscience，1960 年。

[5] 人的感觉力和想象力是否真的可能已经进化为适于处理人体大小尺度的环

境，而非适于尺度上要小十亿倍的微观世界呢？

[6] 马拉·贝勒，《量子对话：革命的成功》（*Quantum Dialogue: The Making of a Revolution*，芝加哥：芝加哥大学出版社，1999 年），第 22 页。

[7] Patrick A. Heelan，《量子力学与客观性》（*Quantum Mechanics and Objectivity*，海牙：尼约夫出版社，1965 年），第 23 页。

[8] 波恩，《我的生活和观点》（纽约：Scribner's，1968 年），第 216 页。

[9] 海森堡，《物理和物理以外：邂逅与对话》（*Physics and Beyond: Encounters and Conversations*，纽约：Harper and Row，1971 年），第 60 页。

[10] 同上，第 61 页。

[11] 海森堡，"运动学和力学关系的量子力学重新解释"（On the Quantum-Mechanical Reinterpretation of Kinematic and Mechanical Relations），《物理学杂志》33（1925），第 879～893 页；B. L. van der Waerden，《量子力学的来源》（*Sources of Quantum Mechanics*，阿姆斯特丹：北荷兰，1967 年），第 261 页。

[12] 海森堡，*AHQP* 访谈，1963 年 2 月 15 日。

[13] "关于运动学和力学关系的量子解释"（Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen），刊于《物理学杂志》53（1925），第 893 页。

[14] 引自贝勒的《量子对话》，第 43 页。

[15] 波恩，"诺贝尔奖晚宴发言"（Remarks at Le Banquet Nobel），1954 年诺贝尔文学奖（斯德哥尔摩：Royale P. A. Norstedt & Stoner，1955 年）。

[16] 波恩，《我这一代的物理学》（*Physics in My Generation*，纽约：培格曼出版公司，1969 年），第 100 页。

[17] Nancy Greenspan，《确定性世界的终结：波恩的一生与科学》（*The End of the Certain World: The Life and Science of Max Born*，纽约：基本图书公司，2005 年），第 127 页。

[18] 发表前，有件事情令他们有点震惊。这件事情就是剑桥大学一位叫做狄拉克的学生所发表的一篇论文。狄拉克在剑桥拿到了海森堡论文的副本，对它进行了研究，得到了与波恩和约尔当一样的结论（只是描述略有不同）。同时，狄拉克还听到了海森堡的报告，应用新的记号，指出了 ***q*** 数与 ***p*** 数之间的不同。这些变量不是经典变量，即不满足对易定律。狄拉克不久后将这些经典变量称为 c 数，它们指的是量子力学变量（q 数）。

[19] 引自派斯的《内界》（*Inward Bound*，纽约：牛津大学出版社，1986 年），第

258 页。

[20] 詹摩尔指出的一点颇具讽刺意味：这种尝试把类似太阳系的图景从原子物理中去除的做法，其数学基础正是从天文学家用来计算行星轨道的方法中推导出来的所谓特征方程。

[21] 引自贝勒的《量子力学解释的起源：1925—1927》，第 81 页。

[22] 薛定谔，“从微观力学到宏观力学的连续过渡”（The Continuous Transition from Micro- to Macro-Mechanics），《自然科学》（*Die Naturwissenschaften*）28（1926），第 664～666 页，载于薛定谔的《论文集》，第 45～61 页。

[23] 詹摩尔，《概念的发展》（The Conceptual Development），第 271 页。

[24] 贝勒，《量子力学解释的起源：1925—1927》，第 85、89 页。

[25] 波恩致 E. Schrödinger 的信，1926 年 11 月 6 日，刊于 *AHQP*。

[26] 薛定谔，《论文集》，第 45-61 页。

[27] 贝勒，《量子力学解释的起源：1925—1927》，第 93 页。

[28] Cassidy，《不确定性》（*Uncertainty*），第 215 页。

[29] 薛定谔，《论文集》，第 46 页。

[30] 同上，第 59 页。

[31] 薛定谔致维恩的信，1926 年 6 月，引自 Cassidy，《不确定性》，第 214 页。

[32] 贝勒，《量子力学解释的起源：1925—1927》，第 207 页。

[33] 贝勒，《量子对话》，第 410 页。

[34] 海森堡，《物理和物理之外》，第 73 页。

[35] 泡利把该解释放在了一篇论文的脚注里：“关于气体退简并与顺磁性”（Über Gasentartung und Paramagnetismus），*Zfp*41，1927

[36] 引自贝勒，《量子力学解释的起源：1925—1927》，第 137 页。

[37] 泡利致海森堡的信，1926 年 10 月 19 日，见 A. Hermann、K. Meyenn 和 V. Weisskopf，《沃尔夫冈·泡利：玻尔、爱因斯坦和海森堡之间的科学通信》（*Wolfgang Pauli: Wissentschaftlicher Briefwechsel mit Bohr, Einstein, Heisenberg, u. a.*，纽约：施普林格出版社，1979 年），第 347 页。

[38] 海森堡致泡利的信，1926 年 11 月 15 日，同上，第 355 页。

[39] 海森堡致泡利的信，1926 年 11 月 23 日，同上，第 359 页。

[40] 实际上，约尔当的文章含蓄地表述了许多人在初次看到不确定性原理时不禁会想到的：电子和其他微小物体的确有位置和动量，人们无法捕捉到这两个量是因为测量仪器有缺陷。这种缺陷甚至可能是无法避免和消除的。

[41] John H. Marburger, III，“海森堡不确定关系的历史渊源是有问题的”（A

Historical Derivation of Heisenberg's Uncertainty Relation is Flawed），《美国物理学杂志》（*American Journal of Physics*）76（2008），第 586～587 页。

[42] 贝勒，《量子力学解释的起源：1925—1927》，第 217 页。

[43] 出处同上，第 318 页。

[44] 海森堡，*AHQP* 访谈，1963 年 2 月 25 日。

[45] 贝勒，《量子力学解释的起源：1925—1927》，第 245 页之后。

[46] 玻尔，《原子论与自然描绘》（*Atomic Theory and the Description of Nature*，剑桥：剑桥大学出版社，1934 年），第 54 页。

[47] 引自"玻尔的哲学"（The Philosophy of Niels Bohr），作者 Aage Peterson，刊于《原子科学家通报》（Bulletin of Atomic Scientists）19，no. 7（1963）。

[48] 引自贝勒，《量子力学解释的起源：1925—1927》，第 248 页

[49] P. W. Bridgman，"科学新视界"（The New Vision of Science），*Harper's*，1929 年 3 月，第 443～451 页。

[50] 我要特别感谢 John H. Marburger, III，是他向我指出了这点。他说："这是一种清晰、富有逻辑、而且具有一致性的考虑互补性问题的方法。它阐明了量子力学是如何通过过去的经典'图像'呈现出来，并与物理学的其他内容漂亮地融合在一起的。这一设想非常清晰，它除去了萦绕在互补性上的神秘色彩。所有发生的一切就好比是格式塔转换（Gestalt-switch），即从用经典观点看待微观世界到接受希尔伯特空间图像的艰难过程，经典概念也由此自然形成。在该转变过程中，玻尔发挥了桥梁的作用。"

后　记

找到回家的路

我们是可以帮陌生人找到回家的路的，而且可以把他们准确无误地带回家。

——斯蒂芬·邓恩（Stephen Dunn），
Walking Light: Memoirs and Essays on Poetry

我把通向方程的路途叫做旅程，不过这个比喻是有误导性的。因为它暗示了通向稳妥的、既定的目的地的过程是平稳顺利的。其实，大部分方程只是人们理解过程的顶峰。这一过程本身并不是平坦的，旅行者常会因为到达的地方与自己想象的目的地不同而激动不已。上述比喻也在错误地暗示，旅行者仅仅是憧憬大自然的旁观者。事实上，旅行者是积极的参与者，他们通过了解自身与大自然之间的相互影响的变化，相应地改变了自己的思想。

但是这一比喻的确也有它好的一面，它抓住了这样一点：随着近景象的出现和其他景象的逝去，以及整体的图景围绕着新路标的重新组织，旅行中的每一步是如何对旅行者的看法进行重新调整的。随着新世界的展现，旅行者认为重要的事物也在发生微妙的变化。有些变化并非是因为特定的一个事件——某个特征、某个发现、某项技术或者某个人——而是因为旅程本身。这就是哲学家所谓的人类行为的历史真实性（historicity）。每一群旅行者都继承了一幅图景、一种思维方式、与此伴随的种种不满足感以及消除这些不满足感所需努力的方向。在最终的旅途中，整个图景就发生了转变，并固定下来。在该过程的每一步中，世

界似乎都具有非凡的多样性，世界的秩序似乎并不如它与生俱来展现给人们的那般。人们在大自然中发现的秩序乃是源于之前的探索和旅程。不过，世界的这种多样性也使得它具有另一个内在秩序的线索，通过调查研究，就能够看清。赫维赛德对麦克斯韦工作的评价也可用于本书提到的所有科学家："只有改变了表述形式[①]，我才能看得更清楚。"这些科学家不满足于所见，具有对事物真相的洞察力，也有能力组织实施调查研究查出真相（哲学家将这一过程称为"诠释循环"）。这一旅程永远没有尽头，每一个新的发现都会对整个图景产生影响。至于处于不断变化之中的实际背景、仪器条件和理论背景的影响就更不必提了。人永远不会满足，永远不会停止期待，也永远不会不再去组织调查研究。只有通过这种方式，才会出现科学，否则科学就会一文不值，或者根本不可能出现。

不过，大部分时候，人们关心的还是方程和方程的用处，而不是这些方程是如何发现的。人们一般习惯于去关注直接观察到的世界的那一部分。这不难理解，而且这么做也有充分的理由。但是，研究科学共同体或者某个科学家发现方程的过程，我们也能学到很多——从无知到科学的过程。

第一，我们会看到方程发现的历程是非常不同的。本书描述的有些历程很简短，几分钟的时间就能掌握。毕达哥拉斯定理就是一个例子。没有受过任何数学训练的人不仅能够掌握它，而且还能体验发现的乐趣。其他一些历程就比较长了：人们发现 $F=ma$ 和 $F_g=Gm_1m_2/r_2$ 可以说是用了几百年，甚至是几千年的时间。有些历程中科学家们相互之间不断通信，才最终导致了方程的发现，如 $E=mc^2$、热力学第二定律和不确定性原理等。还有一些历程几乎是由一人完成的，如爱因斯坦的引力的广义相对论方程和薛定谔的波动方程。不过这些人即便是单独工作的，实际上也

① 这里指电磁方程的表述形式。——译者注

与同事进行了对话。简言之，没有哪项发现能靠某个人完全独立地完成。

我们还发现，方程不仅是简单的科学工具，可以说它们还具有“社会生活”。人们习惯于把方程看成是无声的乐器，只有在科学家和工程师使用它们的时候，它们才会影响世界。然而，方程实际上是活的，可以产生教育甚至是文化上的力量，告诉人们世界是什么，有时还能使人们改变对世界的认识。毕达哥拉斯定理告诉一代代的学生们什么是证明，而牛顿的引力定律则告诉政治思想家们法律的意义。有了热力学第二定律，人们就不会再去空想什么凭空产生的能量；而爱因斯坦的 $E=mc^2$ 和引力的广义相对论方程则从根本上改变了人们对空间和时间的理解。薛定谔方程和海森堡不确定性原理迫使人们重新思考“物体”的意义。

我们从这些旅程中学到的另一点是科学概念的本质。有一种诱人的想法是认为大自然中有一种业已存在的结构，人只是在观察这些结构，并将其“翻译”成数学语言。这些数学语言只是方程和描述，而不是解释或者创新。但是，“翻译”的方式取决于所走过的历程、人们对“翻译”的不满意，以及如何回应人们的不满意。用乔治 •斯坦纳（George Steiner）的话来说就是，我们是在“前进中跌倒”。因此，把科学描绘成仅仅是科学家想出的概念，并对概念进行验证和修正的做法会对人们产生误导。这其中有两个地方是错误的。一个是概念的意义依赖于所有其他概念的意义。概念是牛顿在人们生活其中的鱼缸般世界的一部分，它的意义依赖于鱼缸中的其他的一切。在验证概念时，心里想着自己在验证的就是这个概念，而不是别的其他，就好比是在问牛顿是在波士顿的左边还是右边。你不知道自己在哪里，也不知道其他一切的方位图。我们不光需要玻璃鱼缸中的其余部分，还需要自己从世界中获得的其余经验。我们所相信的科学概念其实是概念加上经验。经验变了（新的实践和技术），世界的概念也就变了。这就是为什么概念不会一成不变，总是在变化或者被修正。此时正确的概念在彼时可能就是不正确的。如果对某个事物没有经验，那么就无法准确地表述它。因此，概念不是决定性的，而是

指示性的。概念需要从人的经验出发，来“指代”某个事物。而且很显然，概念的所指将随着调查研究而发生变化。哲学家把这称为“形式指引”（formal indication）。人们可以对概念进行严格的评估，验证它们是正确的还是不正确的，因此概念只是一个形式，属于封闭系统。概念是指示性的，是说它们指代所有的实验、定义、技术和开放联系等其他事物。而且要保证概念的准确，它们就要依赖于这些事物。这些事物发生变化后，形式元素也能发生变化。实际上，人们是希望它能发生变化的。科学史学家彼得·盖里森（Peter Galison）对此有一段精彩的描述。他写道，以下是理论家的经验：

> 你试着要在某一项上加个减号，但却不行，因为如此一来就破坏了宇称。你又试着向某一项中加入更多的粒子，这也是不允许的，因为现有理论是不能重整的，所以需要无限多个变量才行。于是你又试着从理论中将某个粒子略去，而现在定律出现了无法解释的可能性。你减去一个不同的项，所有的粒子都会消失；把一项分成两项，电荷就不能守恒，而且同时还要满足角动量守恒、线动量守恒、能量守恒、轻子数守恒和重子数守恒。上述限制条件并非是出自基于公理的同一个理论。相反，它们是由一系列实际上相互渗透、介入而产生综合体。像能量守恒等已有 1000 年的历史了。其他限制（如宇称守恒）在被抛弃前也存在了很长一段时间。再有一些，如“自然性”（所有自由变量以整数比例变化）的要求则是起源于最近。它们中有些被研究团队拿来作为几乎不能逾越的障碍，其他一些被推到一边的只是闪着黄色的警示灯。总之，这些限制条件叠加在一起，使人们无法对某些现象加以论断。而其他一些现象（如黑洞）则几乎是无法避免的。[1]

从这些历程中还能看到另一点：科学是一种深刻的情感历程。认识

不到这一点，或者说认为科学体验包含着一部分干巴巴的概念，和另一部分与概念没有关联的情感，这样的人是不理解科学或人类的创造力的。把科学过程分割成概念部分和情感部分不仅可以做到，而且在某些场合也的确是有用的。但是，这毕竟是一个人造的模型，不是真实情况。通过研究这些历程，我们就能深入到抽象层面的背后，看到被隐藏起来的科学的真正运行机制。在这些历程中，我们既能发现不满足感的作用，也能发现好奇、惊愕、困惑和惊讶构筑的小插曲。我们还能看到期待和敏锐之间的不同：有的科学家充满期待，只在结果符合期望时他才会注意到；而另一类科学家却有着敏锐的感觉，他们随时做好了听到意料之外的消息的准备。对于情感，我们不仅能从科学发现的动机中看到，也能从科学家对科学发现所作的回应中看到。对发现的感情上的反应并不是“好，明白了。这是这，那是那。我弄错了，现在知道了”那么简单，而应该是更加微妙的。而且感情上的反应也不光只是对发现而言的。就像利昂·莱德曼（Leon Lederman）所写的：把希望寄托于作出发现，求得功名，“并不是真正的生活”。他接着说：“兴趣和激动一定得是天天都有的——无论在建造新仪器并使其运转的挑战中，在与同事、学生交流的喜悦中，还是在讲座、走廊和期刊上学到新东西的快乐中。”[2]

可是，我们惊奇不光是因为学到了什么，还因为某些意义更加深远的东西。在某些惊奇的时刻，我们瞥见了自身与大自然间的联系，瞥见了大自然的变化不定和人们在大自然中的作用。经验告诉我们：大自然是变化的，所以它可能还是另一番景象。人们也知道，它在将来还是要变化的。此时，我们体会到了爱默生式的高级思索，一种更加微妙的情感降临，令人马上感到新和旧、惊奇和熟悉、若即与若离、神秘和亲近。

在本书即将完成之前不久，我发现竟难以向一位著名的老物理学家描述自己的工作。这位老先生对写给科学家以外人士的科普读物基本上产生不了什么共鸣。对他而言，这都没有什么神奇的。方程一旦掌握了，就好像是很显然、很完整、很合理的，所以难以想象没有方程情况会是

怎样。他按着纯粹职业人士的做法去接近科学，同时也力劝我这么做。他告诉我："如果人们意识到这些方程是如此地显而易见，那么他们就不会对方程感到如此惊奇了。你应该帮助他们。"

我当时真是应该拥抱这位老先生的，他帮我找到了自己一直想要做的。要做的恰恰相反：我应该告诉人们方程并不是显而易见的，让他们找到引领人们发现这些方程的不满足感，还原人们理解方程之后那一刻的惊奇心情。这种惊奇在那一刻同时发现和产生，方程似乎只是我们已有认识的简洁表述。这种认识（如毕达哥拉斯定理）稳固可靠，已然为我们所认识，并被记下来。像上面我认识的那位老先生一类的科学家习惯于关注形式部分（这也是他所说的"显而易见"的意思），而人文领域的哲学家和其他学者则习惯于关注另一部分。对形式部分和人文部分同时加以关注应该是可行的：科学和人文走到一起，病感缺失治愈，吐温作品中巡视水面的老、小飞行员，直盯着图的奴隶小孩和只关注人类生活的苏格拉底，分离出来的形式部分和意义、情感部分又合并到了最初的统一体中。如果能做到这一点，我们就能再次实现理查德·哈里森的孩子发现 1+1=2 的奇妙历程，并且认为方程"不仅对于外部世界的奇妙之处，而且对神和所有人身上的奇妙之处来说都是至关重要的"。这样一个时刻预示着对世界的完全人性化回应（fully human response）的到来。

注　释

[1] Peter Galison，《图像与逻辑：微观物理学的材料文化》（*Image and Logic: A Material Culture of Microphysics*，芝加哥：芝加哥大学出版社，1997 年），第 801 页。

[2] Leon Lederman，"学习的乐趣"（The Pleasure of Learning），《自然》430:5（2004 年 8 月），第 617 页。